

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 25, 2003, 04:01:26 ; Search time 4877 Seconds
(without alignments)
4105.542 Million cell updates/sec

Title: US-09-931-007A-1
Perfect score: 3508
Sequence: 1 MGETLGSPIDPSDSFTDR.....KTEIDMSLHPLLOEIKYKDLX 688

Scoring table:
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgr2_1/USPTO.spool/US09931007/runat_14022003_102553_4678/app.query.fasta_1.839
-DB=GenDb1 -QPM=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=Blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLION=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09931007 @CGN 1.1.3052 @runat_14022003_102553_4678 -MCPu=6 -ICPU=3
-NO_XLIPY -NO_MMAP -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-MARK_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl : *
1: gb_ba : *
2: gb_htg : *
3: gb_in : *
4: gb_cm : *
5: gb_ov : *
6: gb_pat : *
7: gb_ph : *
8: gb_pl : *
9: gb_pr : *
10: gb_ro : *
11: gb_sts : *
12: gb_sy : *
13: gb_un : *
14: gb_vi : *
15: em_da : *
16: em_fun : *
17: em_hum : *
18: em_in : *
19: em_mu : *
20: em_om : *
21: em_or : *
22: em_ov : *
23: em_pat : *
24: em_ph : *
25: em_pl : *
26: em_ro : *
27: em_sts : *
28: em_un : *

29: em_vi : *
30: em_htg_hum : *
31: em_htg_inv : *
32: em_htg_other : *
33: em_htg_mus : *
34: em_htg_pln : *
35: em_htg_rtd : *
36: em_htg_man : *
37: em_htg_vrt : *
38: em_sy : *
39: em_htgo_hum : *
40: em_htgo_mus : *
41: em_htgo_other : *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2231	63.6	1518	9 HS063415	U63415 Human perox
2	2231	63.6	1608	6 AR121467	AR121467 Sequence
3	2231	63.6	1608	9 HSU79012	U79012 Human ligan
4	2217	63.2	1759	9 AF033103	AF033103 Macaca mu
5	2213	63.1	1765	9 AY048695	AY048695 Macaca fa
6	2196.5	62.6	1679	6 AR203332	AR203332 Sequence
7	2196.5	62.6	1679	9 HUMPEPARG	DB3233 Homo sapien
8	2186	62.3	1849	9 AY048698	AY048698 Macaca fa
9	2172.5	61.9	1539	4 AF059245	AF059245 Sus scrofa
10	2171.5	61.9	1758	4 AF103946	AF103946 Sus scrofa
11	2160.5	61.6	1640	4 SSC6757	AJ0066757 Sus scrofa
12	2156	61.5	1782	10 BC021798	BC021798 Mus muscu
13	2153	61.4	1780	10 MM009138	U09138 Mus muscu
14	2148	61.2	1518	10 AB019561	AB019561 Rattus no
15	2148	61.2	1785	10 AF156666	AF156666 Rattus no
16	2143	61.1	1570	10 RNPPARGM2	Y12882 Rattus norv
17	2136.5	60.9	2028	9 AY048697	AY048697 Macaca fa
18	2132	60.8	1815	4 BTPPARG2	Y12420 B. taurus nr
19	2100	59.9	1766	9 HSPPARGM	X90563 H. sapiens m
20	2100	59.9	1808	9 HUMPEPARG	LI0904 Homo sapien
21	2095.5	59.7	1767	9 BC006811	BC006811 Homo sapi
22	2080	59.3	1766	9 AF033343	AF033343 Macaca mu
23	2080	59.3	1794	9 AF033342	AF033342 Macaca mu
24	2076	59.2	1800	9 AY048694	AY048694 Macaca fa
25	2076	59.2	2063	9 AY048699	AY048699 Macaca fa
26	2075	59.2	1713	9 AY048696	AY048696 Macaca fa
27	2074	59.1	1884	9 AY048700	AY048700 Macaca fa
28	2065.5	58.9	1811	6 AX409565	AX409565 Sequence
29	2065.5	58.9	1844	6 AR139020	AR139020 Sequence
30	2065	58.9	2005	6 AR030509	AR030509 Sequence
31	2062	58.8	1636	10 AB011365	AB011365 Rattus no
32	2062	58.8	1798	10 AF246457	AF246457 Rattus no
33	2062	58.8	1838	10 AF246458	AF246458 Rattus no
34	2062	58.8	1848	10 AF156665	AF156665 Rattus no
35	2062	58.8	2130	10 MM001841	U01841 Mus muscu
36	2054	58.6	1428	10 MM010374	U10374 Mus muscu
37	2054	58.6	1709	10 CGPPARGRN	Z30972 C. griseus m
38	2053.5	58.5	1630	4 SSC6756	AJ006756 Sus scrofa
39	2047	58.4	1759	4 OC084893	U84893 Oryctolagus
40	2045	58.3	1428	10 AF317514	AF317514 Cavia por
41	2045	58.3	1711	4 BTPPARG1	Y12419 B. taurus nr
42	2035	58.0	1800	10 U01664	U01664 Mus muscu
43	2028	57.8	1428	5 AF484684	AF484684 Anas plat
44	1933	55.1	2089	5 AB045597	AB045597 Gallus ga
45	1933	55.1	2089	5 AF163811	AF163811 Gallus ga

RESULT 1

ALIGNMENTS

HS063415 1518 bp mRNA linear PRI 18-JUL-1996
LOCUS HS063415
DEFINITION Human peroxisome proliferator activated receptor gamma 2 mRNA,
complete cds.
ACCESSION U63415
VERSION U63415.1 GI:1432176
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 1518)
Elbrecht, A., Chen, Y., Cullinan, C.A., Hayes, N., Leibowitz, M.D.,
Moller, D.E. and Berger, J.
TITLE Molecular cloning, expression and characterization of human
peroxisome proliferator activated receptors gamma 1 and gamma 2
JOURNAL Biochem. Biophys. Res. Commun. (1996) In press
REFERENCE 2 (bases 1 to 1518)
Elbrecht, A., Chen, Y., Cullinan, C.A., Hayes, N., Leibowitz, M.D.,
Moller, D.E. and Berger, J.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-1996) Molecular Endocrinology, merck Research
AUTHORS Labs., P.O. Box 2000, Mail Drop R801-265, Rahway, NJ 07060, USA
FEATURES
source Location/Qualifiers
1..1518
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="fat cell"
1..1518
/codon_start=1
/product="peroxisome proliferator activated receptor gamma
2"
/protein_id="AAB04028.1"
/db_xref="GI:1432177"
/translation="MGETLGDSPTDPESDSTDTLNLANSCEMTWDTMPFPPNNG
ISVYDLVMEHSHSFDIKPTTVDSSISPHNEDIPETPTDVAAYKYLQEX
QSAIKVEPASPYYSEKTLQNKPHBESNSLMAIECVCDKASGFHYGVHACEGK
GFFRRTRILKLIYRCDLNCRHKSRKACQYCFQCLAGMSHNALRFMRQAEK
EKLAETISSDIDOLNPESADRLAKHLHYDSIKSPFTTKAAVILIGKTDSPV
IYDNNSLMEGDKIKFKHPILOESKEVARIRCOGQREVEAORTETAKSIPG
VNDLNOVYTLTKYGVHEIITMLASLNKRGVILSEQCGMTREFLSKRPDGG
EPKFEFAKFNALDSDDLAIFLAVIILSDRGLELWKFIEDIIDLQALDLQK
LNHPESQLEFKLLQKMTDLNROIYTEHVOLLQVTKTETDMSLPLQEIYKDY"

BASE COUNT 438 a 363 c 346 g 371 t
ORIGIN

Alignment Scores:
Pred. No.: 1.32e-159 Length: 1518
Score: 2231.00 Matches: 443
Percent Similarity: 87.72% Conservative: 0
Best Local Similarity: 87.72% Mismatches: 0
Query Match: 63.60% Indels: 62
DB: 9 Gaps: 1

US-09-931-007a-1 (1-688) x HS063415 (1-1518)

OY 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20
|||||
Db 1 ATGGGTGAACCTCTGGAGATTCTCTATTGACCCAGAAAGGATTCCTTCTCATATACA 60
OY 21 LeuSerAlaAsnIleSerGlnIleMetThrMetValAspThrGluMetProPheThrPro 40
|||||
Db 61 CTGTCTGCAAAACATATCAACAAGATGACCATGTTGACACAGAGATGCATCTGGGCC 120
OY 41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
|||||
Db 121 ACCAACCCTTGGATCAGCTCCGATCTCTCCGTAATGGAAGACACCTCCACCTCTTT 180
OY 61 AspIleLeuProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80
|||||
Db 181 GATATCAAGCCCTTACTACTGTTGACTTCTCCACATTTCTACTCCACATTACGAGAAC 240
OY 81 IleProPheThrArgThrAspProValAlaAspTyrTyrTyrAspLeuIleGln 100

|||||
Db 241 ATTCATTACAAAGAACATCCAGTGGTTGCAGATTACAAATGATAGACTGAACCTCA 300
OY 101 GltTyrGlnSerAlaIleIleYsValGluProAlaSerProProTyrTyrSerGluYsThr 120
|||||
Db 301 GAGTACCAAGAGCAATCAATCAAAAGTGAGCCTGCATCTCCACCTTATTAATTCAGAAACT 360
OY 121 GlnLeuTyrAsn----- 124
|||||
Db 361 CAGCTCTACAAATTAACCTCATATGAAGCCTTCCAACTCCCTCATGGCAATTGAATGTGCT 420
OY 124 ----- 124
Db 421 GTGTGTGAGATAAAGCTTCTGATTTTCATGTGAGATTGATGTTGGAAGATGCAAG 480
OY 124 ----- 124
Db 481 GGTTCCTCCGAGAACATCAATGAGATTGAACTTATCTATGACAGATGATCTTAATCTGT 540
OY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
|||||
Db 541 CGGATCCACAAAAAAGTAGAATATAATGTCAGTACTGTGCTTCCAGAAATGCCCTTGCA 600
OY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluYsGluYs 158
|||||
Db 601 GTGGGATGTCTCATTAATGCCATCAGTTTGGCGGATGCCAGCGCAGAGAGAGAG 660
OY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
|||||
Db 661 CTGTTGGGAGATCTCCAGTATATCCAGCAGCTGAAATCCAGATCCGCTGCCTCCGG 720
OY 179 AlaLeuAlaYsHisLeuTyrAspSerTyrIleYsSerPheProLeuThrLysAlaYs 198
|||||
Db 721 GCCCTGGCAAAACATTTGTATGACTCATATCAATAAAGTCTTCCCTGACCAAAAGCAAG 780
OY 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218
|||||
Db 781 GCGAGGCGGATCTTGACAGGAAGAACACAGCAAAATCCATTCGTTATCATGATCATG 840
OY 219 AsnSerLeuMetMetGlyGluAspLysIleYsPheYsHisIleThrProLeuGlnGlu 238
|||||
Db 841 AATTCCTTAATGATGGAGAAATAAATCAAGTTCAAAACATCAACCCCTCGAGGAG 900
OY 239 GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258
|||||
Db 901 CAGACCAAAAGAGTGGCCATCCGATCTTTCAGGGCTGCGCAATTCGCTCCGTGAGGCT 960
OY 259 ValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278
|||||
Db 961 GTGCAGAGATCAGACAGATATGCCAAAGCATTTCTGGTTTGTAAATCTGACTTGAAC 1020
OY 279 AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
|||||
Db 1021 GACCAAGTAACCTCTCCAAATATGAGATCCACGAGATCATTTACACATGCTGGCCTCC 1080
OY 299 LeuMetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGluPhe 318
|||||
Db 1081 TTGATGATTAAGATGGGTTCTCATATCCGAGGCCCAAGCCTTCATACAAAGGAGTTT 1140
OY 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338
|||||
Db 1141 CTAAAGAGCCTGCGAAACCTTTTGTGACTTATGAGAGCCCAAGTTGAGTTCTGTG 1200
OY 339 LysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle 358
|||||
Db 1201 AAGTTCAATGACCTGGAATATGATGACAGCGACTTGGCAAAATTAATATGCTGCTTAATT 1260
OY 359 LeuSerGlyAspArgProGluLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
|||||
Db 1261 CTCAGTGGAGACCGCCAGGTTTGTGTAATGTGAAGCCCATTTGAAGACATTCAAGACAAC 1320
OY 379 LeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPhe 398
|||||

Db 1321 CTGCTACAAAGCCCTGGAGCTCCAGCTGAAGCTGAACCCCTGAGCTCTACAGCTGTTT 1380
Qy 399 AAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeu 418
Db 1381 GCCAGCTGCTCCAGAAATGACAGACCTCAGACAGATTGTGACGAGACACGCTGACGCTA 1440
Qy 419 LeuGlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeuGlnIle 438
Db 1441 CTCGACGTCGATCAGAAAGACGAGACAGACATGAGTCTTCACCCGCTCTCGACGAGATC 1500
Qy 439 TyrLysAspLeuTyr 443
Db 1501 TACAAAGACTGTAC 1515

RESULT 2
AR121467
LOCUS AR121467 1608 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 3 from patent US 6159734.
ACCESSION AR121467
VERSION AR121467.1 GI:14105043
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 1608)
McKay,R., Borchers,A.H. and Baker,B.F.
TITLE Antisense modulation of peroxisome proliferator-activated receptor
gamma expression
JOURNAL Patent: US 6159734-A 3 12-DEC-2000;
FEATURES
source Location/Qualifiers
1..1608
/organism="unknown"
BASE COUNT 462 a 380 c 361 g 405 t
ORIGIN

Alignment Scores:
Pred. No.: 1,416-159 Length: 1608
Score: 2231.00 Matches: 443
Percent Similarity: 87.72% Conservative: 0
Best Local Similarity: 87.72% Mismatches: 0
Query Match: 63.60% Indels: 62
DB: Gaps: 1

US-09-931-007A-1 (1-688) x AR121467 (1-1608)

Qy 1 MetGlyGlnThrLeuGlySperProIleAspProGluSerAspSerPheThrAspThr 20
Db 91 ATGGGTGAACCTCTGGAGATTCTCTATTGACCCAGAAAGGATTCCTTCACTGATACA 150
Qy 21 LeuSerAlaAsnIleSerGlnLumetThrMetValAspThrGluMetProPheThrPro 40
Db 151 CAGTCGCAACATATCACAAGAAATGACCATGTTGACACAGAGATGCCATTCTGGCCC 210
Qy 41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisPhe 60
Db 211 ACCAACTTGGGATGAGCTCCGCTGATCTCTCGTAATGAGAGACACCTCCACCTCTT 270
Qy 61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80
Db 271 GATATCAACCCCTCAGCTACTGTGACTTCTCCACATTCTACATTCACATTACGAGAC 330
Qy 81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGln 100
Db 331 ATTCATTACACAAGAAAGATCCAGTGGTTCAGATTAACAATAGACTGAAACTTCAA 390
Qy 101 GluTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThr 120
Db 391 GAGTACCAAGAGCAATCAAAAGTGGAGCTGCATCTCCACCTTATTATCTGAGAAAGACT 450
Qy 121 GlnLeuTyrAsn 124
Db 451 CAGCTCTACAAATAGCCTCATGAGAGCCTTCCAACTCCCTCATGCGCAATTGAATGTCTG 510

Qy 124 ----- 124
Db 511 GTCGTGGAGATAAGCTTTCGATTTCACATATGAGATTCAATGCTTGTGAAGATGCAAG 570
Qy 124 ----- 124
Db 571 GGTTCCTCCGGAACAATCAGATTGAAGCTTATCTATGACAGATGTGATCTTAATCTGT 630
Qy 125 ----- ArgAsnLysCysGlnTyrCysAspPheGlnLysCysLeuAla 138
Db 631 CGGATCCACAAAAAAGTGAATATAATGTCAAGTACTGTGGTTTCAGAAATGCCCTTGCA 690
Qy 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158
Db 691 CTGGGGATGTCTCATTAATGCCATCAGTTTGGCGGATGCCACAGCCGAGAGAGAGAAAG 750
Qy 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
Db 751 CTGTTGGGGAGATCTCCAGTATATCCAGCTGAATCCAGAGTCCGCTGACCTCCGG 810
Qy 179 AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198
Db 811 GCCCTGGCAAAACATTTGATGACTCATATCAATCAAGTCCCTCCGTAACCAAGCAAG 870
Qy 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218
Db 871 GCGAGGCGGATCTTACAGAGAAAGACAAACAGCAATCACCATTGCTTATCTATACATG 930
Qy 219 AsnSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGlu 238
Db 931 AATTCCTTAATGATGGAGAGATAAATCAAGTTCAAACATCATCACCCCTCGCAGAG 990
Qy 239 GlnSerLysGluValAlaIleArgIlePheGlnGlnCysGlnPheArgSerValGluAla 258
Db 991 CAGAGCAAGAGGTGGCCATCCGATCTTTCAGGGCTCCCATGTTCCGCTCGAGAGCT 1050
Qy 259 ValGlnGluIleThrGlyTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278
Db 1051 GTCGACGAGATCAGACAGATGCCAAAGACATCTCTGTTTGTAACTTGACTTGAAAC 1110
Qy 279 AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
Db 1111 GACCAAGTAACTCTCTCAAAATATGAGTCCACAGATCATTTACACAAATCTGCTGCC 1170
Qy 299 LeuMetAsnLysAspGlyValLeuIleSerGlnGlnGlnGlyPheMetThrArgGluPhe 318
Db 1171 TTGATGAATTAAGATGGGTTCTCTATCCAGGGCCCAAGGCTTCATGACAAGGAGGATTT 1230
Qy 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338
Db 1231 CTAAGAGCCCTCGGAAGCCTTTTGTGCTGACTTATGAGAGCCCAATTTGAGTTGCTGTG 1290
Qy 339 LysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle 358
Db 1291 AAGTTCAATGACACTGGAATTTGATGACAGCGACTTGGCAATTTATTTGCTCTATTATT 1350
Qy 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
Db 1351 CTCAGTGGAGACCGCCGAGTTTCTGAATGTGAAGCCCATTTGAAGACATTCACAAACAC 1410
Qy 379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPhe 398
Db 1411 CTGCTACAAAGCCCTGGAGCTCCAGCTGAGCTGAAGCAACCCCTGATCTCTCAGCTGTTT 1470
Qy 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeu 418
Db 1471 GCCAAGCTGCTCCAGAAATGACAGACCTCAGACAGATTGTTCAGAGGAGACAGTGCAGCTA 1530
Qy 419 LeuGlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeuGlnGluIle 438
Db 1531 CTCGAGGTGATCAAGAAAGACGAGACAGATGAGTCTTCCACCGCTCTCGAGAGATTC 1590
Qy 439 TyrLysAspLeuTyr 443

Db 1591 TACAAAGACTGTAC 1605

RESULT 3
HSU79012

DEFINITION Human ligand activated transcription factor PPARgamma2 mRNA,
complete cds.

ACCESSION U79012

VERSION 079012.1 GI:1711116

KEYWORDS

SOURCE Homo sapiens.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 1608)
Mukherjee, R., Jow, L., Croston, G.E. and Paterniti, J.R. Jr.
Identification, characterization, and tissue distribution of human
peroxisome proliferator-activated receptor (PPAR) isoforms
PPARgamma2 versus PPARgamma1 and activation with retinoid X
receptor agonists and antagonists

JOURNAL J. Biol. Chem. 272 (12), 8071-8076 (1997)

MEDLINE 97218249

PUBMED 9065481

REFERENCE 2 (bases 1 to 1608)
Mukherjee, R., Jow, L., Croston, G.E. and Paterniti, J.R. Jr.
Direct Submission
Submitted (20-NOV-1996) Cardiovascular Research, Ligand
Pharmaceuticals, Inc., 9393 Towne Centre Drive, San Diego, CA
92121, USA

FEATURES

source Location/Qualifiers

1..1608
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="heart"
91..1608
/note="nuclear receptor"
/product="ligand activated transcription factor
PPARgamma2"
/protein_id="AAC51248.1"
/db_xref="GI:1711117"

CDS

1/translacion= "WGETLGDSPIDESDSFTDLSANI SOEMTWDT EMPFPTNG
ISSVLSVMEHSHSDIKPPTVDSSISPHYDIDIPTRDPVAAYKYDLKQEY
OSAIKVEPASPYYSEKTOYLNK PHEEPSNLMAIECRVCGKASGFGVYACGCK
GEFRITIKLIYDRCDLNCRIHKS RNKCOCRKOCIAVMSNNAIRFGMPQAEK
EKLAIESSDIDOLNPESADLRALAKHLYDSIKSFPLTKAKARAILGKTKDSPV
IYDMSILMGEDEKIKFKHTTLPLOQESKEVALRIFGCGCFRSVEAOETTERAKSIDGF
VNLIDNDVTLTKTGVEHIIYTMFLASLNKQDVLISEQGEWTFREFLSLRPFQDFM
EPKFEFAVFNALIEDSDLAIFIAVILISGDRGLNVKPIEDIDNMLQALEQLK
LNHPSSOLFALQKMTDLROIYVEHVLQVLIKTEEDMSLHPLQELVKDLY"

BASE COUNT 462 a 380 c 361 g 405 t

ORIGIN

Alignment Scores:

Pred. No.: 1..41e-159 Length: 1608

Score: 2231.00 Matches: 443

Percent Similarity: 87.72% Conservativeness: 0

Best Local Similarity: 87.72% Mismatches: 0

Query Match: 63.60% Indels: 62

Db: 9 Gaps: 1

US-09-931-007a-1 (1-688) x HSU79012 (1-1608)

QY 1 MetGlyGluThrLeuGlyGlySerProIleAspProGluSerAspSerPheThrAspThr 20

Db 91 ATGGGGAAACTCTGGGAACTCTCTATTGACCCAGAAAGGATTCCTTCACTGATACA 150

QY 21 LeuSerAlaAsnIleSerGlnIleuThrMetValAspThrGluMetProPheThrPro 40

Db 151 CTGTCTGCAACATATACACAAATGACCATGTTGACACAGAGATGCCATTCCTGGCC 210

QY 41 ThrAsnPhgGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60

Db 211 ACCAACTTGGATCAGCTCCGGATCTCTCCGTAATGAGACACATCCACTCTT 270

QY 61 AspIleIysProPheThrThrValAspPheSerSerIleSerThrProHisItyrGluAsp 80

Db 271 GATATCAACCCCTTCACTCTCTGCTTGTGACTTTCAGACTTTCACATTCACAGAAC 330

QY 81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuIleuGln 100

Db 331 ATTCATTCACAGAACATCAGATGCTGTCAGATTACAAATATGACCTGAACCTTCA 390

QY 101 GlnTyrGlnSerAlaIleIysValGluProAlaSerProProTyrTyrSerGluIysThr 120

Db 391 GAGTACCAAGTCATCAAG 450

QY 121 GlnLeuTyrAsn 124

Db 451 CAGCTCTACAAATAGCCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 510

QY 124 124

Db 511 GTCTGTGAGATAAGCTTCTGATTTCACTATGAGATTCAATGCTTGTGAGAGATGCAAG 570

QY 124 124

Db 571 GGTTCCTCCGAGAACATCAGATTGAAGCTTATCTATGACAGATGATGCTTAAGTGT 630

QY 125 -----ArgAsnIysCysGlnTyrCysArgPheGlnIysCysLeuAla 138

Db 631 CGATCCACAAAAAAGTGAATATATGTCAGTGTGCTGCTTGCAGAAATGCTTGTGA 690

QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluIysGluIys 158

Db 691 GTGGGATGTCATATATATCCATCAGATTGGGCGGATCCACAGCCGAGAGAGAGAG 750

QY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178

Db 751 CTGTTGGCGAGATCTCCAGTATATGACAGCTGAATCCAGAGAGAGAGAGAGAGAG 810

QY 179 AlaLeuAlaIleHisLeuTyrAspSerTyrIleIysSerPheProLeuThrIysAlaIys 198

Db 811 GCCCTGGCAAAACATTTGATGACATCATTAAGTCTTCCCTGACCAACCAAG 870

QY 199 AlaArgAlaIleLeuThrcIyLysThrTrAspIysSerProPheValIleTyrAspMet 218

Db 871 GCGAGGCGCATTTGACAGAGAAAGACAAACAGCAATCCATTCGTATATCATCATG 930

QY 219 AsnSerLeuMetMetGlyGluAspIysIleIysPheIysHisIleThrProLeuGlnGlu 238

Db 931 AATTCCTTAATGATGAGAGAAATAAATCAAGTTCAAACATCATCACCCCTGCAGAG 990

QY 239 GlnSerIysGluValAlaIleArgIlePheGlnIyCysGlnPheArgSerValGluAla 258

Db 991 CAGAGCAAAAGAGGTGGCCATCCGATCTTTCAGGGCTGCACATTTGCTCCCTGAGAGCT 1050

QY 259 ValGlnGluIleThrcIyLysSerIleProGlyPheValAsnLeuAspLeuAsn 278

Db 1051 GTGCAGAGATATCACAGATATGCCAAACATTCCTCGTTTGTAAATCTTACATTGAAC 1110

QY 279 AspGlnValThrLeuLeuIysTyrGlyAlaHisGluIleIleTyrThrMetLeuAlaSer 298

Db 1111 GACCAAGTAACCTCTCTCAATATGAGATCCACAGATCATTTACACAACTGCTGCCTCC 1170

QY 299 LeuMetAsnIysAspGlyValLeuIleSerGlnIyGlnIyPheMetThrArgGluPhe 318

Db 1171 TTGATGAATAAAGATGGGTTCTCATATCCGAGGGCCCAAGGCTTATGCAAGGAGATT 1230

QY 319 LeuIysSerLeuArgIysProPheGlyAspPheMetGluProIysPheGluPheAlaVal 338

Db 1231 CTAAAGAGCCTCGAAAGGCTTTTGGTGTACTTATATGAGAGCCCAAGTTTGAAGTTGCTGTG 1290

QY 339 IysPheAsnAlaIleGluLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle 358

Db 1291 AAGTCATGACCTGATAGATGACAGCAGCTTGCAAAATTTATTGCTGTCATTAATT 1350
 QY LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
 Db 1351 CTCATGTGGAGACCGCCGAGTTTGTGATGTGAAGCCCATTTGAACATTCACACAAAC 1410
 QY LeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPhe 398
 Db 1411 CTGCTACAAAGCCCTGAGCTTCAGCTGAAGCAGCAACCCCTGAGTCTCCACAGCTGTTT 1470
 QY AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeu 418
 Db 1471 GCCAAGCTGCTCCAGAAATGACAGACCTCAGACAGATTGTCACGGAACACCTGAGCTTA 1530
 QY LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnIle 438
 Db 1531 CTGCAAGGATGATTAAGAACGAGACGACATGATCTTCCACCCCTCTGACAGAGATC 1590
 QY TyrLysAspLeuTyr 443
 Db 1591 TACAGAGACTGTAC 1605
 RESULT 4
 AF033103 1759 bp mRNA linear PRI 27-NOV-1997
 LOCUS Macaca mulatta peroxisome proliferator-activated receptor gamma 2
 DEFINITION (PPARG2) mRNA, complete cds.
 ACCESSION AF033103
 VERSION AF033103.1 GI:2645730
 KEYWORDS
 SOURCE Macaca mulatta
 ORGANISM Macaca mulatta
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 Cercopithecinae; Macaca.
 REFERENCE 1 (bases 1 to 1759)
 AUTHORS Hotta,K., Gustafson,T.A., Yoshioka,S., Ortmeier,H.K., Bodkin,N.L.
 and Hansen,B.C.
 TITLE Relationships of PPARG and PPARG2 mRNA levels to obesity, diabetes,
 and hyperinsulinemia in rhesus monkeys
 JOURNAL 2 (bases 1 to 1759)
 REFERENCE Hotta,K., Gustafson,T.A., Yoshioka,S., Ortmeier,H.K., Bodkin,N.L.
 and Hansen,B.C.
 TITLE Direct Submission
 JOURNAL Submitted (05-NOV-1997) Department of Physiology, University of
 Maryland, School of Medicine, 10 South Pine Street, MSTF Room 6-00,
 Baltimore, MD 21201, USA
 FEATURES
 source
 1. 1759
 /organism="Macaca mulatta"
 /db_xref="taxon:9544"
 /tissue_type="adipose"
 1. 1759
 /gene="PPARG2"
 25. 1542
 /gene="PPARG2"
 /note="transcription factor"
 /codon_start=1
 /product="peroxisome proliferator-activated receptor gamma
 2"
 /protein_id="AAB87480.1"
 /db_xref="GI:2645731"
 /translation="MGETIGDSPIDPESDFTDLSANISOETIMVDTMPMPPTNG
 ISSYDLSVMDHSHSPDIKPTPTVDPSSISAPHEDI.PFTPDMAVDKXLDLQEX
 OSAIKVEPASPPIYSEKTLQYNKHPHPSNSLAIETRCVGDRAKSGHYHACGCK
 GFEFRTRKLTIYRCDLNCRIHKRSNKCQYCFQCCLAVGMSHNAIRGRMQAK
 EKLAEISSDQNPESADLRALAKHLIYDSYIKSFPLTAKARAILTGTDTKSPV
 IYDMSLMVGEDEKIKFKHPILEOSKEVAIRIFOGCFREVAOETETAKSIPEF
 VNLIDNVLTGKYGVHIIYTMASLMNKGVIISGCGFMEPEFLSKRKPGDM
 EPRFEAVKFALELDDSDLAIFAVIILSDRGLNVRIDEIONLQALELQK
 LNHPESSQLEFKLLQKMTDLKQIVTEHVQLQVTKTETDMSLHPLEQETKDY"
 BASE COUNT 532 a 407 c 376 g 444 t

ORIGIN
 Alignment Scores:
 Pred. No.: 1,8e-158 Length: 1759
 Score: 2217.00 Matches: 440
 Percent Similarity: 87.08% Conservative: 5
 Best Local Similarity: 86.11% Mismatches: 4
 Query Match: 63.20% Indels: 62
 DB: 9 Gaps: 1
 US-09-931-007a-1 (1-688) x AF033103 (1-1759)
 QY 1 MetGlyGlnThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20
 Db 25 ATGGGTAAACTGCTGGAGATTCCTCTTATGACCAGAAAGCATTCCTCAGTATGACA 84
 QY 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrPro 40
 Db 85 CTGCTCGCAACATATTCACAGAAATTAACATGTTGACACAGAGATGCCATTCGGCCC 144
 QY 41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
 Db 145 ACCAATTGGATCAGCTCCGTGATCTCTCCGTAATGACGACCACTCCACTCTTT 204
 QY 61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisIleTyrGluAsp 80
 Db 205 GACATCAAGCCCTTCACACGCTTCTGCTTCCAGCATTTCGTCCCATTTATGAAAC 264
 QY 81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGln 100
 Db 265 ATTCATTCACAGAACAGATCCGATGCTCGGATTAAGTATGACCTGAACCTTCAA 324
 QY 101 GlnTyrGlnSerAlaIleLysValGlnProIleAspProPheThrProTyrTyrSerGlnLysThr 120
 Db 325 GAGTACCAAGTGCATCAATCAAGTGGAGCTGCTTCCACCTTATTTATTCGAGAGACT 384
 QY 121 GlnLeuTyrAsn 124
 Db 385 CAGCTCTCAATTAAGCCTCAGAGAGCCTTCCAACTCCCTCATGGAATGCAATGTCGT 444
 QY 124 124
 Db 445 GTCTCGGAGATAAGCTTGTGATTCACATGAGATTGCTTGTGAAGATGCAAG 504
 QY 124 124
 Db 505 GGTTCCTCCGAGACATCAGATTGAGCTTATCTATGACAGATGTCATTAACTGT 564
 QY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
 Db 565 CGGATCCACAAAAGAAATTAATGTCAGTACGTGCGTTTCAAAATGCGCTTGCA 624
 QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlnTyrGlyMetProGlnAlaGluLysGluLys 158
 Db 625 GTGGGAGTGTCTCAATTAAGCCATCAGGTTGGGGAGCCACAAAGCCGAGAGAGAG 684
 QY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
 Db 685 CTGTTGGCGGAGATCTCCAGTATATGACACAGCTGAAACCCAGAGTCCGCTGACCTCCGG 744
 QY 179 AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198
 Db 745 GCCCTGGCAAAACATTTGTATGACATCATACATAAAGTCCCTCCGCTGACCAAGCAAG 804
 QY 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218
 Db 805 GCGAGGGGATCTTGACAGAGAAAGCAACAAATACCATTCGTTATATGACATG 864
 QY 219 AsnSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGlu 238
 Db 865 AATTCCTTAATGATGGAGAAATTAATCAAGTTCAACACATCACCCCACTGACAGAG 924
 QY 239 GlnSerLysGluValAlaIleArgIlePheGlnGlnLysGlnPheArgSerValGluAla 258

|||||
Db 925 CAGACGAAGAGGTGGCCATCCGATCTTTCAGAGGCTCGACTTCGCTCCGGAGGCT 984
OY 259 VALGInGluIleThrGluTyrAlaLysSerIleProGlyPheValAlaLeuAsnLysPhe 278
Db 985 GGGCAGGAGATCACAGATGATGCCAAAGCATTCCTGGTTGCTGAATCTTGACCTTGAAC 1044
OY 279 AAGGInValThrLeuLeuTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
Db 1045 GACCAAGTAACTCTCTCAATATATGAGTCCATGATCATTTTACCATGCTGGCTCC 1104
OY 299 LeuMetAsnLysAspGlyValLeuIleSerGluGlyGlyPheMetThrArgLysPhe 318
Db 1105 TTGATGAAATAAAGATGGGCTTCTCATATCCGAGGGCCAGGCTTCATGACAAGGGATTT 1164
OY 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338
Db 1165 CTAAAGAGCTCGGAAGGCTTTGGTGACTTATGAGACCCCAAGTTGAGTTTGCTGTG 1224
OY 339 LysPheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIleIle 358
Db 1225 AAGTTCAATGCATGGAATATGATGACAGCAGCTTGGCAATTTATTCCTGCTCATTTAT 1284
OY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
Db 1285 CTCAGTGGAGATCGCCAGCTTGTGATGTGAAGCCCATGGAACATTCACAAGCAAC 1344
OY 379 LeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGlnSerSerGlnLeuPhe 398
Db 1345 CTGCTACAAAGCCCTGGAGCTCCAGCTGAAGCTGAACCCCTGAGTCCGCGAGCTGTTT 1404
OY 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeu 418
Db 1405 GCCAAGCTCTCCAGAAATGACAGACCTCAGACAGATTGTCACGGAACACGTGACGCTA 1464
OY 419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnLeuIle 438
Db 1465 CCGCAGGTGATCAGAAAGCGAGAGACATGAGCTTCACCCGCTCTCGCAGAGGATC 1524
OY 439 TyrLysAspLeuTyrAlaThrPalaIleLeuThr 449
Db 1525 TACAAAGACTTGTACTAGCAGAGAGTCTTGAGC 1557

RESULT 5
LOCUS AY048695 1765 bp mRNA linear PRI 21-MAY-2002
DEFINITION Macaca fascicularis peroxisome proliferator-activated receptor
gamma 2 (PPARGgamma2) mRNA, complete cds.
ACCESSION AY048695
VERSION AY048695.1 GI:15723731
KEYWORDS
SOURCE
ORGANISM
Macaca fascicularis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecidae; Macaca.
REFERENCE
AUTHORS Zhou, J., Wilson, K.M. and Medh, J.D.
TITLE Genetic analysis of four novel peroxisome proliferator activated
receptor gamma splice variants in monkey macrophages
JOURNAL Biochem. Biophys. Res. Commun. 293 (1), 274-283 (2002)
MEDLINE 12054596
PUBMED 22050014
TITL 2 (bases 1 to 1765)
AUTHORS Zhou, J., Wilson, K.M. and Medh, J.D.
JOURNAL Direct Submission
200 Submitted (27-JUL-2001) Internal Medicine, The University of Iowa,
200 Hawkins Drive, Iowa, IA 52242, USA
FEATRES
source
1. 1765
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/cell_type="macrophage"

gene 1. 1765
/gene="PPARGgamma2"
25. 1542
CDS /gene="PPARGgamma2"
/note="transcription factor"
/codon_start=1
/product="peroxisome proliferator-activated receptor gamma
2"
/protein_id="AA05262.1"
/db_xref="GI:15723732"
/translation="MGETLGDSPIDPESDSTDTLISANISQETIMVDTENPFPNTNG
ISVDLSVMDHSHFDIKPPTVDFESSISAPHYEDIPETKDPVADVKYDLQLEY
QSAIKVEPASPYPYSEKTOLYNKHHEEPSNLSIEGVCGRKASGFHGVHACEGK
GFPRRTIKLIVRCDLNCRIHKSRNKCQYCRFOCLAVGSHNAIRFGMRQAEK
EKLAEISSDIDOLNPESADLRALAKHLIDYGIKSPFLTKAKARALITGKTDKSPY
IYDMSLSIMGEDIKPKHIIPILOQSKFEVARIRGCGOFSEVAVOETEVAKSIFPK
VNLNDQVTLKATGVHEIITMLASLMNKDGVITSGOGFMTREFLSKLRPGDPM
EPKFEFAVKFNALEDDSDLAIFLAVIILISDRGLNVPRIEDIQNLQALEQLK
LNHPSSOLFVKLLQKMTDLROIYTEHVOLLQVYIKKTEMDSLHPLQEIYKDLX"
BASE COUNT 537 a 409 c 376 g 443 t
ORIGIN
Alignment Scores:
Pred. No.: 3,61e-158 Length: 1765
Score: 2213.00 Matches: 439
Percent Similarity: 86.89% Conservative: 5
Best Local Similarity: 85.91% Mismatches: 5
Query Match: 63.08% Indels: 62
D: 9 Gaps: 1
US-09-931-007a-1 (1-688) x AY048695 (1-1765)
OY 1 MetGlyGluThrLeuGluYAspSerProIleAspProGluSerAspSerPheThrAspThr 20
Db 25 ATGGGTGAACACTCTGGAGATTCCTCTATTGACCCAGAAAGGATTCCTTCACTATACA 84
OY 21 LeuSerAlaAsnIleSerGlnLumethrMetValAspThrGluMetProPheThrPro 40
Db 85 CTGCTGCAAAACATATCACAGAATAATTCAGATGTTGACACAGAGATGCATTTGGGCC 144
OY 41 ThrAsnPhlegIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
Db 145 ACCAAGCTTTGGAGTACGCTCCGATGATCTCTGTAATGAGAGACCACTCCACATCCCTT 204
OY 61 AspIleLysProPheThrThrValAspPheSerIleSerThrProHisTyrGlnLysP 80
Db 205 GACATCAACCCCTTCACTACTGCTTCTCAGACATTTCTGCTCCACATTTATATAAC 264
OY 81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGln 100
Db 265 ATTCATTCACAAAGAACAGATCCGATGGTGGGATTTAAGTAGACCTGAACTTCAA 324
OY 101 GluTyrGlnSerAlaIleLysValIleProAlaSerProProTyrTyrSerGluLysThr 120
Db 325 GGTGACCAAGAGCAATCAAGATGAGAGCTGCTTCCACCTTATATTCTGAGAAAGACT 384
OY 121 GlnLeuTyrAsn 124
Db 385 CAGCTCTCAATTAACCTCAGAGAGACCTTCCAACTCCCTCATGGCAATGCATGTGCT 444
OY 124 124
Db 445 GTCGCGGAGATAAGCTTCTGGATTCCACTATGAGATTCATGCTTGGAAGATGCAAG 504
OY 124 124
Db 505 GGTTCCTCCGAGAACATCAGATTGAAGCTTATCTATGACAGATGTGATCTTAACGT 564
OY 125 ArgAsnLysCysGlnTyrCysAspPheGlnLysCysLeuAla 138
Db 565 CGGATCCACAAAAAGTAGAATAATATGTCAGTACTGCTGCTTTCACAAATGCTTCA 624
OY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158

|||||
Db 625 GTGGGATGCTCATATATGATCCATGAGTTGGGCGGATGCCAAGCCGAGAGAGAG 684
QY 159 LeuLeuAlaGluIleSerSerAspIleAspGluLeuAsnProGluSerAlaAspLeuArg 178
Db 685 CTGTGGGGGAGATCTCCAGTATGATGACAGCAGTGAACCCAGATCCGCTGACCTCCGG 744
QY 179 .AlaLeuAlaLysHisLeuThrAspSerTyrIleLeuSerPheProLeuThrLysAlaLys 198
Db 745 GCCCTGGCAAAACATTTGTATGACTCATACATTAAGTCTTCCCTCCGCTGACCAAGCAAG 804
QY 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218
Db 805 GCGAGGCGGATCTTACAGGAAGACACAGCAATACCATTCGTATGTATGACATG 864
QY 219 AsnSerLeuMetMetGlyGluAspLysIleLeuSphelyshHisIleThrProLeuGlnIle 238
Db 865 AATTCTTAATGATGGAGAGATAAATCAAGTTCAAAACATCACCCCTGACGAG 924
QY 239 GlnSerLysGluValAlaIleArgIlePheGlnGlyGlnPheArgSerValGluAla 258
Db 925 CAGACCAAGAGAGTGTATCCGCAATCTTTCAGGCGTCCAGTTCCGCTCCGTGGAGCT 984
QY 259 ValGlnIleIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278
Db 985 GTGCGAGAGATCACAGAGATATGCCAAAGCAATCCGTGTTCTGTAATCTTGACTTGAAC 1044
QY 279 AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
Db 1045 GACCAAGTACTCTCTCAATATGAGAGTCCATGAGATCATTTACATATCTGCCCCC 1104
QY 299 LeuMetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGluPhe 318
Db 1105 TTGATGATTAAGATGAGGGGTTTTCATATCCAGGCGCAAGCTTCATACAGGAATTT 1164
QY 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338
Db 1165 CTAAAGACCTGCGAAAGCCTTTTGTGACTTATGAGCCCAAGTTGAGTTGCTGTG 1224
QY 339 LysPheAsnAlaLeuGlnLeuAspAspSerSerAspLeuAlaIlePheIleAlaValIle 358
Db 1225 AAGTTCAATGCACTGGAAATGTGATGACAGCATTTGGCAAAATTTTGTGCTCATTTAT 1284
QY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGluAspAsn 378
Db 1285 CTCATGAGAGATCGCCAGGTTTGTGATGTGAAGCCCATTTGAAGACATTCAGACACAC 1344
QY 379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPhe 398
Db 1345 CTGCTACAGCCCTGAGCTCCAGCTCAAGCTGAACACCTCGAGTCTCGCAGCTGTT 1404
QY 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeu 418
Db 1405 GCCAAGCTGCTCCAGAAATATGACAGCTTCAGACAGATTTGACGGAACAGCTGAGCTTA 1464
QY 419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnIle 438
Db 1465 CTGCGAGTGCATCAGAAACAGAGACACATGAGTCTTCAACCCGCTCTCGAGAGATC 1524
QY 439 TyrLysAspLeuTyrAlaThrAlaIleLeuThr 449
Db 1525 TACAAGGACTGTACTAGACAGAGATCCTGAGC 1557

RESULT 6
AR203332 1679 bp DNA linear PAT 20-JUN-2002
LOCUS AR203332 1679 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 5 from patent US 6365361.
ACCESSION AR203332
VERSION AR203332.1 GI:2149694
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 1679)
AUTHORS Taniuchi,T. and Mizukami,J.
TITLE Method for identifying or screening agonist and antagonist to PPAR
JOURNAL Patent: US 6365361-A 5 02-Apr-2002;
FEATURES
source location/Qualifiers
1. 1679
BASE COUNT 477 a 405 c 377 g 420 t
ORIGIN
Alignment Scores:
Pred. No.: 5,98e-157 Length: 1679
Score: 2196.50 Matches: 439
Percent Similarity: 87.15% Conservative: 2
Best Local Similarity: 86.76% Mismatches: 2
Query Match: 62.61% Indels: 63
DB: Gaps: 2
US-09-931-007a-1 (1-688) x AR203332 (1-1679)
QY 1 MetGlyuThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20
Db 159 ATGGGTGAACCTGGGAGATTCCTATTGACCCAGAAACGATTCTCTCATGATACA 218
QY 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrPro 40
Db 219 CTGTCTGCAAAACATATCAACAGAAATGACCATGGTTGACACAGAGATTCGCTGCCC 278
QY 41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
Db 279 ACCAATCTTGGATCAGCTCCGTGATCTCCGTAATGAGAAACACTCCCATCTCTT 338
QY 61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80
Db 339 GATATCAAGCCCTTCACTACGTGTTGACTTCCAGCATTTCTACTCCACATTAAGAAAGAC 398
QY 81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGln 100
Db 399 ATTCCATTACAAACAGATCCAGTGTTCAGATTTACAAAGTATGACATCGTAACCTTCAA 458
QY 101 GluTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThr 120
Db 459 GAGTACCAAGAGTCAATCAAGTGAAGCTGATCTCATCTCACTTATTTATTCGAGAAAGCT 518
QY 121 GlnLeuTyrAsn 124
Db 519 CAGCTTACATTAAGCTCATGAAGAGCTTCCAACTCCCTCATGCAATTGAATGTCGT 578
QY 124 124
Db 579 GTCGTGAGATTAAGCTTCGTGATTTCACTATGAGAGTTCATGCTTGTGAAGATGCAG 638
QY 124 124
Db 639 GGTTCCTCCGGAACAATCAGATTGAGCTTATCATGACAGATGTGATCTTAACGT 698
QY 125 125
Db 699 CGGATCCACAAAAAAGTAAATTAATGTCAGTACGTGCTGTTTCAGAAATGGCTTCA 758
QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158
Db 759 GTGGGATGTCTCATATGCGCATGAGTTTGGCGGATCCGACAGGCCGGAAGAGAAAG 818
QY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
Db 819 CTGTTGGCGGAGATCTCCAGTATATGACACAGCTCAATCCAGATCCGCTGACCTCGT 878
QY 179 ---AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAla 197
Db 879 CAGGCCCTGCGCAAAACATTTGTATGACTCATACATTAAGTCTTCCGCTGACCAAAAGCA 938
QY 198 LysAlaArgAlaIleLeuThrGlnGlyThrThrAspLysSerProPheValIleTyrAsp 217

```
|||||
Db 939 AAGGAGAGGCATCTTGCACAGAAAGACACAGCAATCCATTCCTATCATGAC 998
Oy 218 Metasensleuemetciylgluasplysileysphelyshiilethrprouleuin 237
Db 999 ATGATTCCTTATATGATGGAGACATTAATCAAGTTAAACACATCCACCCCTGCAG 1058
Oy 238 gluinserlysgluvalalaileargilpehnginlycysgluphearseryalgu 257
Db 1059 GAGCAGAGCAAGAGAGTGGCCATCCGATCTTCAAGGGCTGCAGTTGCTCCGGGAG 1118
Oy 258 AlaValGIngluilethrgluTyralalysSerileproglyPheValAsnleuaspleu 277
Db 1119 GGTGGCAGAGATCATCAGAGATGCCAAAAGCATTCCTGGTTTGTAAATCTTGACTTG 1178
Oy 278 AsnAspGInValThrLeuLeuTyrglyValHisgluileileTyThrMetLeuAla 297
Db 1179 AACGACCAAGTACTCTCTCAATATATGGATGCCAGATCATTTACACATGCTGGCC 1238
Oy 298 SerleuMetAsnlyAspDylValLeuileSerGluGInglyPheMetThrArglu 317
Db 1239 TCCCTGATGAATAAAGATGGGGTCTCATATCCGAGGGCCAAAGGCTTCATGACAAAGGAG 1298
Oy 318 PheLeuTySerLeuArglySerProheglyAspPheMetGluProLysPheGluPheAla 337
Db 1299 TTTCTAAAGAGCCTCGAAAGCCTTTGGTGAATTTATGGACCCCAAGTTTGAGTTTGCT 1358
Oy 338 VallyspheAsnAlaLeuGluLeuAspAspSerAspLeuAlailePheleleValIle 357
Db 1359 GTGAAGTTCAATGCATGGAATTAATGACAGCGACTGGCAATATTTATTTGCTGTCATTT 1418
Oy 358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProileGluAspIleGlnAsp 377
Db 1419 ATTCTACATGGAGACCCGCCAGAGTTTGGTGAATGTGAAGCCCATTTGAAGACATTCAGAC 1478
Oy 378 AsnLeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerGlnLeu 397
Db 1479 AACCTGCTACAAAGCCCTGGAGCTCCAGCTGAGCTGAAGCCATCTGAGCCCTCACAGCTG 1538
Oy 398 PheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGln 417
Db 1539 TTTGCCAAGCTCTCTCAGAAATGACAGACCTTCAGACAAATGTGTACGGAAACAGCTGCG 1598
Oy 418 LeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnlu 437
Db 1599 CTACTGACAGGTATCAGAAAGACGAGACAGCATGAGTCTCACCCGCTCTCTGACGAGG 1658
Oy 438 IleTyryLysAspLeuTyrr 443
Db 1659 ATCTACAAGACTTGCTAC 1676

RESULT 7
HUMPPARG 1679 bp mRNA linear PRI 06-FEB-1999
LOCUS Homo sapiens mRNA for PPAR gamma2, complete cds.
DEFINITION Homo sapiens mRNA for PPAR gamma2, complete cds.
ACCESSION D83233.1 GI:1199768
VERSION D83233.1 GI:1199768
KEYWORDS PPAR gamma2.
SOURCE Homo sapiens adipose cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 1679)
AUTHORS Kato, S.
TITLE PPAR gamma2
JOURNAL Unpublished
AUTHORS Kato, S.
REFERENCE 2 (bases 1 to 1679)
JOURNAL Unpublished
TITLE Direct Submission
AUTHORS Submitted (29-JAN-1996) Shigeaki Kato, Tokyo University of
JOURNAL Agriculture, Department of Agricultural Chemistry, Sakuragaoka
1-1-1, Setagaya-Ku, Tokyo 156, Japan (E-mail:ktakeem.u-tokyo.ac.jp,
Tel:03-5477-2318, Fax:03-5477-2619)
```

```
COMMENT On or before Mar 17, 1999 this sequence version replaced
          gi:2077947, gi:1197807.
FEATURES
          Submitted (18-Jan-1996).
          Location/Qualifiers
            source
              1. 1679
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /tissue_type="adipose"
                159..1679
                /codon_start=1
                /product="PPAR gamma2"
                /protein_id="BAI1894.1"
                /db_xref="GI:1945622"
                /translation="MGETLDDSPIDPSDSFTDPLSANISOEMTMDVETIAFPPTNG
                ISSVDLSVMDHSHSPDIKPTTVDFSSISPTPHEDIPTRTDVPVADYKYLKLEY
                OSAIKVPSAPPYSEKTOLYNKHPSNSLMAIECRVCGAKSGFVHACGECK
                GFPRRTRIKTIYRCDLNCRHKHSRKNKOYCFQKCLAVGNSHNAIRGRIAOAKR
                EKLATISSDIDQNPESADIRQALNHLNIDSYIKSPPLTKAKARAILLTKTTDSFP
                VTYDMSLMGEDKIKFRHITPRLDQSKREVALRIEQGCFRSEVALEVEAKSIIPG
                FVNLDLDQVTLKYGVHEIITPMLASLMKMDGLISEGQFMTREFLKLRFPGDP
                MEKFERAVKFNLELDDSLAIPFAVILISGDRGLLNKPIEDIDNLLQALEDL
                KLNHRESSOLFAPKLLQKMTDLROIIVTEHVQLLOYIKKTTEDMSLHPLDEIYDLY"
BASE COUNT 477 a 405 c 377 g 420 t
ORIGIN
Alignment Scores:
Pred. No.: 5,98e-157 Length: 1679
Score: 2196.50 Matches: 439
Percent Similarity: 87.15% Conservative: 2
Best Local Similarity: 86.76% Mismatches: 2
Query Match: 62.61% Indels: 63
DB: Gaps: 2
US-09-931-007a-1 (1-688) x HUMPPARG (1-1679)
Oy 1 MetGlyLutThrLeuGlyAspSerProIleasProGluSerAspSerPheThrAspThr 20
Db 159 ATGGGTCAAACTCTGGGAGATTCCTCTATGACACAGAAAGCATTCCTTCACATGATACA 218
Oy 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrPro 40
Db 219 CTGTGTCAAACTATATCAAGAAATGACAAAGTGTGACACAGAGATGCAATTCGTGGCCC 278
Oy 41 ThrAspPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
Db 279 ACAAATTTGGGATAGCTCCGTGGATCTCTCCGTAATGAAGAACAATCCACATCCTTT 338
Oy 61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisTyrgluAsp 80
Db 339 GATATCAAGCCCTTCACACTACTGTGACTTCTCCAGCATTTCTACTCCACATTAAGAGAC 398
Oy 81 IleProPheThrArgThrAspProValAlaAspTyryLysTyryAspLeuLysLeuGln 100
Db 399 ATTCATTTCACAGACAGACAGATCCAGTGGTTCAGATTACAGATGACCTTAACCTTCAA 458
Oy 101 GluTyrgInserAlaIleLysValGluProAlaSerProProTyryTySerGluLysThr 120
Db 459 GAGTACCAAGAGCAATCAAGAGTGGAGCTGCATCTCCACCTTATTTATTTGAGAGAAGCT 518
Oy 121 GlnLeuTyrrAsn----- 124
Db 519 CAGCTCTCAATTAAGCCTCATGAAGAGCCTTCCAACCTCCCTCATGGCAATTGAATGTGCT 578
Oy 124 ----- 124
Db 579 GTCGTGGAGATAAAGCTTTCGATTTTCATATGAGAGTTTCATGCTTGTGAAGATGCAAG 638
Oy 124 ----- 124
Db 639 GGTTCCTCCGAGAACACATGAGATTTATCTATGACAGATGATGATCTTAAGTCT 698
Oy 125 -----ArgAsnLysCysGlnTyrrCysArgPheGlnLysCysLeuAla 138
          |||||||
```

```

Db      699 CGGATCCACAAAAAGTAGAAATAATGTCAGTCTGCGTTTCAGAAATGCTTGCA 758
Qy      139 VALGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGluAlaGluIleGly 158
      |||
Db      759 GTGGGATGTCATATATCCATCAGGTTTGGGCGGATGCGACAGGCCGAGAGAGANG 818
Qy      159 LeuLeuAlaGluIleSerSerAspIleAspGluLeuAsnProGluSerAlaAspLeuArg 178
      |||
Db      819 CTTGTGGCGAGATGTCCTCCAGTATGACAGCACTCAATCCAGAGTCCGCTGACCTCGCT 878
Qy      179 ---AlaLeuAlaLysHisLeuIleArgPheSerPheIleLysSerPheProLeuThrLysAla 197
      |||
Db      879 CAGGCCCTGGCAAAACATTGTTGATGACTCATATAAATCCCTCCGCTGACCAAGCA 938
Qy      198 LysAlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAsp 217
      |||
Db      939 AAGCGAGGCGCATCTTGACAGGAAAGACACAGACAAATCCCATTCCTTATCTATGAC 998
Qy      218 MetAsnSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGln 237
      |||
Db      999 ATGAATTCCTTATGATGGAGAGAGATAAATCAAGTCAACACATCACCCCGCTGACG 1058
Qy      238 GluGlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGlu 257
      |||
Db      1059 GAGCAGAGCAAGAGAGGCGCCATCCGATCTTTCAGGGCTGCCAGTTGCTCCGTGGAG 1118
Qy      258 AlaValGlnGluIleThrGlyLysThrLysSerIleProGlyPheValAsnLeuAspLeu 277
      |||
Db      1119 GCTGTGCAGAGATCACAGATGATGCAAAAGCATTCCTGGTTTGTAAATCTTGACCTTG 1178
Qy      278 AsnAspGlnValIleThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAla 297
      |||
Db      1179 AAGCAGCAAGTAACTCTCTCAATATGAGATGACAGCATATTCATTACCATGCTGGCC 1238
Qy      298 SerLeuMetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGlu 317
      |||
Db      1239 TCCCTGATGATAAAGAGTGGGCTCTCATATCCGAGGGCCAAAGGCTCTCATGCAAGGGG 1298
Qy      318 PheLeuIleSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAla 337
      |||
Db      1299 TTTCTAAAGAGCTCGGAAGGCTTTTGGTGTGACTTATGAGAGCCCAAGTTTGAAGTTTGC 1358
Qy      338 ValLysPheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIle 357
      |||
Db      1359 GTGAAGTTCATGACTGGAATTAGATGACAGCACTGGCAATATTATTCCTGTCAT 1418
Qy      358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAsp 377
      |||
Db      1419 ATTCTCAGTGGAGACCGCCAGGTTTGTGAAATGTAACCCCATTTGAAACATTCAGAGAC 1478
Qy      378 AsnLeuLeuGluAlaLeuGluLeuAsnHisProGluSerSerGlnLeu 397
      |||
Db      1479 AACCTGCTCAAGCCCTGAGCTCCAGCTGAAGCTGAACCATCTGATGCTTCACAGCTG 1538
Qy      398 PheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValIleThrGluHisValGln 417
      |||
Db      1539 TTTCGCAACCTGCTCCAGAAATGACAGACCTCAGACAGATTGTCACGGAACAGCTGGCAG 1598
Qy      418 LeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuGlnGln 437
      |||
Db      1599 CTACTGCGAGTATCAGAAAGAGAGACATGATGCTTCACCGCTCTCTGACAGAG 1658
Qy      438 IleTyrLysAspLeuTyr 443
      |||
Db      1659 ATCTACAGAGACTTGATC 1676

```

```

RESULT 8
LOCUS   AY048698
DEFINITION Macaca fascicularis peroxisome proliferator-activated receptor
gamma 5 (PPARGgamma5) mRNA, complete cds.
ACCESSION AY048698
VERSION 1
GI:21552436

```

```

KEYWORDS
SOURCE
ORGANISM
Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE
1 (bases 1 to 1849)
Zhou,J., Wilson,K.M. and Medh,J.D.
Genetic analysis of four novel peroxisome proliferator activated
receptor gamma splice variants in monkey macrophages
Biochem. Biophys. Res. Commun. 293 (1), 274-283 (2002)
JOURNAL
MEDLINE
2050014
PUBMED
12054596
REFERENCE
2 (bases 1 to 1849)
Zhou,J., Wilson,K.M. and Medh,J.D.
Direct Submission
Submitted (27-JUL-2001) Internal Medicine, The University of Iowa,
200 Hawkins Drive, Iowa, IA 52242, USA
FEATURES
source
1..1849
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/cell_type="macrophage"
1..1849
/gene="PPARGgamma5"
199..1626
/gene="PPARGgamma5"
/note="transcription factor"
/codon_start=1
/product="peroxisome proliferator-activated receptor gamma
5"
/protein_id="AAL05268.1"
/db_xref="GI:21552437"
/translation="MVDTEMPFPTNNGISVDSVWDHSHSFDIPFTTVDFSSIS
APHEIDIPFTTDMVADYKDXDLQRYQNAIKVPSAPYISKTLQYKPHPEPSN
SLMAIEKRCVGDKASGFHYGNVHACGCKGFFRTIRLKLIVRCDLNCRIHKSRNKC
QYCFRQCLAVGMSHNAIRFGMPQAEKRLAEISSDIDLPNEDSADRLAKNHLVD
SYKSPFLTKAKARAILTKTDDKSPFVINDMSLMMGEDKIKRKHITTPQEOSKEVA
IRIQGGQFSGVEAVOITEITPAKSIPEFVINDLNDQVTLKYGVEHETIMMALSMMK
DGVFISRGCPMTPEPLAKSRKPGDPMERKFEPAKVFNNLEDDSLAFIVATLIS
GDRPGLNINVRIEDIDNLDALDELFLQKLHNPSSQLFAKLQKMTDLROIIVTEHVL
LQVIRKETDMSLHPLOEITYKDY"
BASE COUNT 554 a 437 c 396 g 462 t
ORIGIN
Alignment Scores:
Pred. No.: 4,18e-156 Length: 1849
Score: 2186.00 Matches: 438
Percent Similarity: 82.37% Conservative: 6
Best Local Similarity: 81.26% Mismatches: 5
Query Match: 62.31% Indels: 90
DB: 9 Gaps: 2
US-09-931-007a-1 (1-688) x AY048698 (1-1849)
Qy      1 MetGlyGluThrLeuGluIleAspSerProIleAspProGluSerAspPheThrAspThr 20
      |||
Db      25 ATGGGTGAACCTGTGGGAGATTCTCTATTGACCCAGAAAGCATTCCTTCACTGACACA 84
Qy      21 LeuSerAlaAsnIleSerGln-----GluMetThrMetVal 27
      |||
Db      85 CTGCTCGCAACATATCACAAGTAGACAGCTGGGTTTACCATGATGATGATCTCGTACG 144
Qy      28 -----GluMetThrMetVal 32
      |||
Db      145 TCGGTATCCACCACCCCTGGCGCTCCCAAGTGTGGGATTACAGCAATATACCATGATG 204
Qy      33 AspThrGluMetProPheTyrProThrAsnPheGlyIleSerSerValAspLeuSerVal 52
      |||
Db      205 GACACAGAGATGCCATTTTGGGCCACCACTTTGGGATCAGCTCCGATGATCTCTCCGTA 264
Qy      53 MetGluAspHisSerHisSerPheAspIleLysSerProPheThrThrValAspPheSerSer 72
      |||

```

Db 265 ATGAGACACCTCCACCTCTTGACATCAAGCCCTTCACTGTTGACTTCCAGC 324
 QY 115SerThrProHisTyrGluAspIleProPheThrArgThrAspProValAlaAsp 92
 Db 325 ATTTGCGTCCACATTATGAGACATTTCCATTCAAGAACATCGATGGTGGGAT 384
 QY 93 TyrLysTyrAspLeuLysLeuGlnGluTyrGlnSerAlaIleLysValGluProAlaSer 112
 Db 385 TATAGTACACCTGAAACTTCAAGAGTACCAAGTGCATCAAGTGCAGCTGCTTC 444
 QY 113 ProPheTyrSerGluTyrThrGlnLeuTyrAsn----- 124
 Db 445 CCACCTTATTTATCTGAGAGACTCAGCTCTACATTAAGCTTCACGAGAGCCCTTCCAC 504
 QY 124 ----- 124
 Db 505 TCCCATGAGCATCGAATGCTGCTGGGAGATTAAGCTTCTGATTCACATGCA 564
 QY 124 ----- 124
 Db 565 GTTCATGCTTGTGAGAGATGCAAGGCTTCTCCGAGAACATTCAGATTGAAGCTTATC 624
 QY 125 -----ArgAsnLysCysGlnTyr 130
 Db 625 TATGACAGATGTGATCTTAACCTCGGATCCACAAAAAGTGAATAATATGTCAGTAC 684
 QY 131 CysArgPheGlnLysCysLeuAlaValAlaGlyMetSerHisAsnAlaIleArgPheGlyArg 150
 Db 685 TGTGCGTTGAGAAATGCTTGCAGTGGGATGTCTCATATGATCCATCAGTTGGGCGG 744
 QY 151 MetProGlnAlaGluLysGluLysLeuLeuAlaGluIleSerSerAspIleAspGlnLeu 170
 Db 745 ATCCCAACACCCAGAGAGGAGCTGTGGCGAGATCTCCAGTATATGACACGACGCTG 804
 QY 171 AspProGlnSerAlaAspLeuArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLys 190
 Db 805 AACCCAGAGTCCGCTGACCTCCGCGCCCTGCAAAACATTTGATGATCATCATCAATAAG 864
 QY 191 SerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLysThrAspLys 210
 Db 865 TCTTCCCGCTGACCAAGCAAGGAGGCGCATCTTGACAGAAAGAACACAGACAA 924
 QY 211 SerProPheValIleTyrAspMetAsnSerLeuMetCysGluAspLysIleLysPhe 230
 Db 925 TCACCATGCTTATCATGATGAGAAATTCCTTATGATGGAGAAATTAATCAAGTTTC 984
 QY 231 LysHisIleThrProLeuGlnGlnGlnSerLysGluValAlaIleArgIlePheGlnGly 250
 Db 985 AACCAATCCACCCGCTGACGAGAGCAAGAGAGGTGGCTATCCGCATCTTTCAGGGC 1044
 QY 251 CysGlnPheArgSerValGluAlaValAlaGlnGluIleThrGlyTyrAlaLysSerIlePro 270
 Db 1045 TGGCAGTTGCGTCCGAGGAGGCTGTGAGAGATCACAGATGATGCCAAAGCATTTCT 1104
 QY 271 GlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysTyrGlyValHisGln 290
 Db 1105 GGTGTGTAATCTTGACTTGAACGACCAAGTAACTCTCTCAATATGAGAGCCATGAG 1164
 QY 291 IleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyValLeuIleSerGlnGly 310
 Db 1165 ATGATTTACCAATGCTGGCTCTGATGATGAATAAGATGGGTTTTCATATCCGAGGGC 1224
 QY 311 GlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysProPheGlyAspPheMet 330
 Db 1225 CAGGCGTTCATGACAGAGGAATTTCTAAAGAGCCTCGAAAGCCTTGGTGACTTATG 1284
 QY 331 GluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeuAspAspSerAspLeu 350
 Db 1285 GACCCCAAGTTTGAAGTTGCTGGAAGTTCAATGACACTGGAATTAAGATGACACGACTTG 1344
 QY 351 AlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGlyLeuLeuAsnValLys 370
 Db 1345 GCATATTTATTTCTGCTGCTATTTCTCATGTGAGATGCGCCAGGTTTGGTGAATGTGAAG 1404

QY 371 ProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGlnLeuGlnLysLeuAsn 390
 Db 1405 CCCATTGAAGACATTCAGACACACGCTTACAAAGCCCTGGAGCTCCAGCTGAAGCTGAAC 1464
 QY 391 HisProGlnSerSerGlnLeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArgGln 410
 Db 1465 CACCTGAGTCCCGCAGCTGTTTGGCAAGCTGCTCCACAAATGACACACCTCAACAG 1524
 QY 411 IleValThrGluHisValGlnLeuLeuGlnValIleLysLysThrGlnThrAspMetSer 430
 Db 1525 ATGTGACGAGACACGCTGAGCTACTGCAAGTGATCAAGAGACGAGACATGAGCT 1584
 QY 431 LeuHisProLeuLeuGlnGluIleTyrLysAspLeuTyrAlaThrAlaIleLeuThr 449
 Db 1585 CTTCACCCGCTCCTGACAGAGATCTCAAGAGACTTGTACTAGCAGAGAGCTCCGAGC 1641
 RESULT 9
 AF059245 1758 bp mRNA linear NAM 21-APR-1998
 LOCUS AF059245
 DEFINITION Sus scrofa peroxisome proliferator-activated receptor gamma 2
 (PPARG) mRNA, complete cds.
 ACCESSION AF059245
 VERSION AF059245.1 GI:3065956
 KEYWORDS
 SOURCE Sus scrofa.
 ORGANISM Sus scrofa.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE 1 (bases 1 to 1758)
 AUTHORS Houseknecht, K.L., Bidwell, C.A., Portocarrero, C.P., Spurlock, M.E.
 and Heady, A.
 TITLE Cloning and regulation of porcine peroxisome proliferator-activated
 receptor gamma
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1758)
 AUTHORS Houseknecht, K.L., Bidwell, C.A., Portocarrero, C.P., Spurlock, M.E.
 and Heady, A.
 TITLE Direct Submission
 JOURNAL Submitted (13-APR-1998) Animal Sciences, Purdue University, 1026
 Poultry Science Building, West Lafayette, IN 47906-1026, USA
 FEATURES
 source location/Qualifiers
 1..1758
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 1..1758
 /gene="PPARG"
 17..1531
 /gene="PPARG"
 /codon_start=1
 /product="peroxisome proliferator-activated receptor gamma
 2"
 /protein_id="AAC1348.1"
 /db_xref="GI:3065957"
 /translation="MGERLQDSLDPESDARFDLTLSNISQEVTVNVDTEMPMPNFGI
 SSVDLSVMDHSHSFEDIKPFYTVDFSSSTFHYEDIPRPADPMADYKRYLQIDYQ
 SAIKVEPSPYSEKLTQLYNKPHPEPNSLMAITCRVCGKASGFHVGAGCGCG
 FERRIRLKLTYDRCDLNCRIHKSRNKCVCYCRQKCLAVGMSHNAIFEGMPAREK
 KILATISSDIOLNPESADIRALAKHIDSYIKSPRLKAKARALITGKTDKSPFY
 YDMSLKMGEIKRKHTPLQESKREVAIRFQGCQPSRYEAVOETETAKNIPGV
 NIDLNDQVTLKRYGVHETITVTLASLKMKDVLLISEGGFMTREFLKLRFQDPMF
 PKFEFVAENMLDDSDLAIFIAVILISGRPGILNVKPIEDIODNLQALELQKL
 NHPESQCFKALKQKMTPLROIETHEVOLLQVIKKTEIDMSLHPLQETRYDLY"
 BASE COUNT 522 a 416 c 390 g 430 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4,1e-155 Length: 1758
 Score: 2172.50 Matches: 440
 Percent Similarity: 82.58% Conservative: 15
 Best Local Similarity: 79.85% Mismatches: 25
 Query Match: 61.93% Indels: 71
 DB: 4 Gaps: 4

US-09-931-007a-1 (1-668) x AF059245 (1-1758)

QY 1 MetGlyGluThrLeuGlyAspSerProIleaspProGluSerAspSerPheThrAspThr 20
|||||
Db 17 AAGGGGAAACTCTGGAGATTCTCTTATTTAGCCAGAAAGCGATGCCCTTC--GACAG 73

QY 21 LeuSerAlaAsnIleSerGlnIleMetThrMetValAspThrGlnMetProPheThrPro 40
|||||
Db 74 CTGCTGGCAACATTTCACAAAGGGTGACCATGTTGACCCGGAATGCCGTTTGGCCC 133

QY 41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
|||||
Db 134 ACCAACTTGGATCAGCTCTGTGACCTGTGGTGATGAGACACACCTCCACTCCTTT 193

QY 61 AspIleIysProPheThrThrValAspPheSerIleSerThrProHisTyrGluAsp 80
|||||
Db 194 GACATCAACCCCTTCCACCTGTGATTTCTCCACATTTCACACTTCACACTATCAAGAC 253

QY 81 IleProPheThrArgThrAspProValAlaAspTyrIleTyrAspLeuIleGln 100
|||||
Db 254 ATCCATTTCCGAGAGCTGATCCATGTTGACAGATTATAGTACGACCTGAGCTCCAG 313

QY 101 GlnTyrGlnSerAlaIleIysValGluProAlaSerProProTyrTyrSerGluIleThr 120
:::
Db 314 GACTACCAAGTGCATCAAGTGAAGCTGTGTCCCGCTTATTTCTGAAAAAGACT 373

QY 121 GlnLeuTyrAsn----- 124
|||||

Db 374 CAGCTGTACAAACAACTCAGAGAGCCCTTCCAACTCCCTCATGCAATTGATGCCGT 433

QY 124----- 124

Db 434 GTCTGTGGGGATAAAGCCTCGGGTTCACATGAGATTGATGCTTGTGAAGATGCAAG 493

QY 124----- 124

Db 494 GGTTCCTTCCGAGAGACTATCAGATTGAAGCTTATTTATGATAGTGGGATTTAACTGT 553

QY 125-----ArgAsnIysCysGlnTyrCysArgPheGlnIleCysLeuAla 138
|||||

Db 554 CGGATCCACAAAAAGATGAATAATGTCAGTACTGTCCGTTTCAGAAAAAGCCCTCT 613

QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGlnIleGlnIle 158
|||||
Db 614 GTGGGATGTCTCATACGCCATCAGGTTTGGCGGATGCCCACAGGCTGAGAGAGAGAG 673

QY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
|||||
Db 674 CTGTGGCAGAGATCTCCAGTATGACGACGACGAGCTGAGACTGTGACTCCGC 733

QY 179 AlaLeuAlaIleHisLeuTyrAspSerTyrIleIysSerPheProLeuThrIleAlaIys 198
|||||
Db 734 GCCGTGGCAAAACCTGTGATGACATCAATAAAGTCCCTCCCTCACCAAAACAAAG 793

QY 199 AlaArgAlaIleLeuThrGlyIleThrThrAspIleSerProPheValIleTyrAspMet 218
|||||
Db 794 GCGAGGCGCATTTTACAGAGAAAGACACAGCAATATCACCTTTGTATCTATATCATG 853

QY 219 AsnSerLeuMetMetClyGluAspIleIysPheIysPheIleThrProLeuGlnIle 238
|||||
Db 854 AATTCCTTAATGATGGAGAAAGATAAATCAAGTCAAAACATCAACCCCTCGAGAG 913

QY 239 GlnSerIysGlnValAlaIleArgIlePheGlnIleGlyCysGlnPheArgSerValGluAla 258
|||||
Db 914 CAGACCAAAAGAGTGGCATTCGATCTTTCAGGGGTGTCACTTTGCTCCCTGAGAGCC 973

QY 259 ValGlnGluIleThrGluTyrAlaIysSerIleProGlyPheValAsnLeuAspLeuAsn 278
|||||
Db 974 GTGCAAGAGATACAGAGATGCCAAGAACATCCTCGGTTTGTAAACCTTGACCTGAAT 1033

QY 279 AspGlnValThrLeuLeuIysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
|||||

Db 1034 GACCAAGTAACCTCCCTAAGATGAGCGTCCACAGATCATTTACACCATGCTGGCTCC 1093

QY 299 LeuMetAsnIysAspGlyValLeuIleSerGlnIleGlnIlePheMetThrArgIlePhe 318
|||||
Db 1094 TTGATGAATTAAGAGAGGGGTCTCATCTCCAGAGGGCCAAAGATTCATATCAAGGAGATT 1153

QY 319 LeuIysSerLeuArgIysProPheGlyAspPheMetGlnProIysPheGluPheAlaVal 338
|||||
Db 1154 CTCAGAGACCTGAGAAACCCCTTTGGTACTTTATGAGGCCCAAGTTGACTTGTCTGTG 1213

QY 339 LysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle 358
|||||
Db 1214 AAGTTCAACGACCTGAGATTTAGATGACAGCGACCTGGCATATTATATGCTGATATT 1273

QY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValIysProIleGluAspIleGlnAspAsn 378
|||||
Db 1274 CTCAGTGGAGACCCGCCAGGTTTGTGAATGTGAAGCCCATTCGAGAGCATTCAGACAAAT 1333

QY 379 LeuLeuGlnAlaLeuGlnLeuGlnLeuIysLeuAsnHisProGluSerSerGlnLeuPhe 398
|||||
Db 1334 TTGCTGCAAGCCTTGAGACTGCTGAGTCAAGTTGAACCAACCCAGAGTCTCCACAGCTTT 1393

QY 399 AlaIysLeuLeuGlnIleMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeu 418
|||||
Db 1394 GCTAAGCTGTCTCCAGAAATGACAGACCTCAGACAGATTGTGACAGACACCTGACCTG 1453

QY 419 LeuGlnValIleIysIysThrGluThrAspMetSerLeuHisProLeuLeuGlnIleIle 438
|||||
Db 1454 CTGCAAGTAATTAAGAAAGACAGAGACGACATGACCTTCCACCACTCTCAGAGAAATA 1513

QY 439 TyrIysAspLeuTyrAlaThrAlaIleLeuThrGlyIleThrThrAspIysSerPro--- 457
|||||

Db 1514 TACAAAGACTGTGAC-----TACCAAAAGATCCACGTTCCACCGCA 1555

QY 458 --PheValIleTyrAspMetAsnSerLeuMetGlyGluAspIysIleIysPheIys 476
|||
Db 1556 ACATCTCCTCTTCTCCATTTGACATATATTGTTGGGGAAAAAATGTGACACTTAA 1615

QY 477 HisIleThrProLeuGlnIleGlnIleSerIysGlu 487
|||||

Db 1616 AAAATTAAGTGTGAAAAAGCGTTTGAAAAAGAA 1648

RESULT 10
AF103946

LOCUS 1539 bp mRNA linear MAM 12-Apr-2000

DEFINITION Sus scrofa peroxisome proliferator activated receptor gamma 2 (PPAR gamma 2) mRNA, complete cds.

ACCESSION AF103946

VERSION AF103946.1 GI:4378476

KEYWORDS

SOURCE

ORGANISM

Sus scrofa.
Sus scrofa.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 1539)
Ding S.T., McNeel R.L. and Mersmann, H.J.
Expression of porcine adipocyte transcripts: tissue distribution and differentiation in vitro and in vivo
Comp. Biochem. Physiol. B, Biochem. Mol. Biol. 123 (3), 307-318 (1999)

MEDLINE 99410873
PUBMED 10481259

REFERENCE 2 (bases 1 to 1539)
Ding, S.T. and Mersmann, H.J.
Direct Submission
Submitted (03-NOV-1998) Pediatrics, Baylor College of Medicine,
1100 Bates St., Houston, TX 77030, USA

FEATURES
source
1..1539
/organism="Sus scrofa"
/db_xref="taxon:9823"
1..1539
/gene="PPAR gamma 2"

gene

CDS

16. .1530
/gene="PPAR gamma 2"
/codon_start=1
/product="peroxisome proliferator activated receptor gamma 2"
/protein_id="MAD19577.1"
/db_xref="GI:4378477"
/translation="MGELGDSLIDPESDAFPTLSANISOEVTVMDTEMPFWPTNGI
SSVDLSVMDHSHSPDIKPFYTVDESSISTPYEDIPEPRADPMVADYKYDLKLDYO
SAIKVEPVSPYSEKTOLYNKPHEPSSSLMAIEKRVCGDASGFVHAGCEGKG
KFRRTIRLKIYDRDLDNCRIRHKSRNKOCYRFOCKLAVGKSHNAREGRPOAEKE
KLARISSDIDLNPSADRLRLAKHLDSYIKSPFLTAARAIITGTTPKSPVY
YMNLSLMWGEDKIKFKHITPLQESKEVALRIFQCGQFSEVAODETSTYANIPREV
NLDELNDQVTLKYGVAHEIITYMLASLMDGVLIISGQFPMREFLSRKPFQDFME
PKFEFAVKFNALIEDSDLAIFAVIILISGDRGLNPKPIEDIODNLQALELOLKL
NHPSOELFAKLQKMTDRLQRIYTEHVHOLLQVYIKKETDMSLHPLLQEIYKDLV"

BASE COUNT 421 a 384 c 369 g 365 t

ORIGIN

Alignment Scores:

Pred. No.:	4.14e-155	Length:	1539
Score:	2171.50	Matches:	432
Percent Similarity:	86.73%	Conservative:	6
Best Local Similarity:	85.54%	Mismatches:	4
Query Match:	61.90%	Indels:	63
DB:	4	Gaps:	2

US-09-931-007A-1 (1-688) x AFI03946 (1-1539)

QY 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThr 20
16 ATGGGGAACCTCTGGAGATCTCTTATTGACCCAGAAAGCGATGCTTC---GACAG 72
21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThr 40
73 CTCTCGCAAACTTTCACAAAGAGGTGACCATGTGACCCAGATGCCGTTTGGCCC 132
41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
133 ACCAACTTGGGATCGAGCTGTGGACCTGTGCGATGAGCAGACCACTCCACTCTTT 192
61 AspIleLeuSerPheThrThrValAspPheSerSerIleSerThrProHisThrGluAsp 80
193 GACATCAAGCCCTTCACACTGTGATTCTCCAGATTTCACACTCCACATCTGAAGAC 252
81 IleProPheThrArgThrAspProValIleAlaAspTyrIleTyrAspLeuGln 100
253 ATCCCATTCGCCGAGAGCTGATCCATGTTGCGAGATTATAGACGACCTCGAAGCTCCAG 312
101 GluTyrGlnSerAlaIleLeuValGluProAlaSerProProTyrTyrSerGluTyrThr 120
313 GACTACCAAGTCCCATCAAAAGTGAGCGCTGTCCCGCTTATTATCTGAAAAGACT 372
121 GlnLeuTyrAsn----- 124
373 CAGCTGTACACAAACCTCAGAGAGCCTCCCACTCCCTCATGGCAATTGAATCCCGT 432
124 ----- 124
433 GTCTGTGGGATAAAGCCTCGGGGTCCTCATGAGATTGATCTTGTGAGAGATGCAG 492
124 ----- 124
493 GGTTCCTCCGAGACTATCAGATTGAAGCTATTATGATAGTGCATCTTAAGTGT 552
125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
553 CGGATCCACAAAAAAGTAAATAATGTCAGTACTGTGCTTGAAGAAATCCCTTGGCT 612
139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluTyrGluTyr 158
613 GTGGGATGTCTCATACGCCATCAGGTTTGGCGGATGCCACAGGCTGAGAGGAGAG 672

QY 159 LeuLeuAlaGluIleSerSerAlaAspGlnLeuAsnProGluSerAlaAspLeuArg 178
673 CTGTGGCAGAGATCTCCAGTGTATTCGACACGACTGAACCCAGACGCTGTGACTCCGC 732
QY 179 AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198
733 GCCCTGGCAAGACCTTGATGACTCATCAATAAATCTTCCCGGTGACCAAAAGCAAG 792
QY 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218
793 GCGAGGCGGATCTTGACAGAGAAAGACCACAGACAAATCACCTTTTCTATGACATG 852
QY 219 AsnSerLeuMetLeuGlyAspLysIleLysPheLysHisIleThrProLeuGlnGlu 238
853 AATTCCTTAATGATGGAGAGAAAGATAAATCAAGTCAACACATCACCCCTGCAGAG 912
QY 239 GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258
913 CAGAGCAAGAGGTGGCCATTCGATCTTTCAGGGGTGTCACTTTCCTCGGAGGCGC 972
QY 259 ValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278
973 GTGCAGAGATCTCAGAGATATGCCAAGACATGCCGTGGGTTTCTAACTTGACCTGAA 1032
QY 279 AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
1033 GACCAAGTAACCTCTCTAAAGTATGGCGTCCACAGATCATTTACACCATGCGGCTCC 1092
QY 299 LeuMetAsnLysAspGlyValIleLeuIleSerGlnGlyGlnGlyPheMetThrArgGluPhe 318
1093 TTGATGAATTAAGACGGGGCTCTCATCTCCGAGGGCAAGATTCATGCAAGGGAGTTT 1152
QY 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338
1153 CTCAGAGCCTGAGAAAGCCCTTGGTGACTTATGAGACCCCAAGTTCAGTTGCTGTG 1212
QY 339 LysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle 358
1213 AAGTTCAACGACACTGGAATTAGATGACAGCAGCACTGGCATATTATTACCTGTCATTATT 1272
QY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
1273 CTCAGTGGAGACCGCCAGCTTGTGAAATGTCAACCCATCGAGGACATTCGAAGCAAT 1332
QY 379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerGlnLeuPhe 398
1333 TTGCTCTCAAGCCTTGGAGCTGACGCTCAAGTTGAACACCCAGAGTCTCCACGCTCTT 1392
QY 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValIleThrGluHisValGlnLeu 418
1393 GCTAAGCTCTCCAGAAATGACAGACCTCAGACAGATTGTGACAGAGCAGCTGCAGCTG 1452
QY 419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIle 438
1453 CTGCAAGTAAATAAGAAAGACAGAGACGACATGAGCCTTACCCACTCTCAAGAGAAATA 1512
QY 439 TyrLysAspLeuTyr 443
1513 TACAAGACTTGTAC 1527

RESULT 11

SSC6757 1640 bp mRNA linear MM 24-SEP-1998

LOCUS SSC6757

DEFINITION Sus scrofa mRNA for peroxisome proliferator-activated receptor gamma 2.

ACCESSION AJ006757

VERSION AJ006757.1 GI:3646469

KEYWORDS peroxisome proliferator-activated receptor gamma 2; PPARg2 gene.

SOURCE Sus scrofa.

ORGANISM Sus scrofa.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

REFERENCE 1 (bases 1 to 1640)

AUTHORS
Grindflek, E.
TITLE
Direct Submission
JOURNAL
Submitted (12-JUN-1998) Grindflek E., Animal Science, Agricultural University of Norway, P.O. Box. 5025, 1432 AS, NORWAY
REFERENCE
2 (bases 1 to 1640)
AUTHORS
Grindflek, E., Sundvold, H., Klungland, H. and Lien, S.
TITLE
Characterisation of porcine peroxisome proliferator-activated receptors gamma 1 and gamma 2: detection of breed and age differences in gene expression
JOURNAL
Biochem. Biophys. Res. Commun. 249 (3), 713-718 (1998)
MEDLINE
98401156
PUBMED
9731203

FEATURES
source
Location/Qualifiers
1..1640
/organism="Sus scrofa"
/db_xref="taxon:9823"
/cell_type="adipocyte"
61..1575
/gene="PPARG2"
61..1575
/gene="PPARG2"
/function="adipocyte differentiation, fat deposition, regulation of adipogenesis"
/codon_start=1
/product="peroxisome proliferator-activated receptor-gamma 2"
/protein_id="CAA07225.1"
/db_xref="GI:3646470"
/db_xref="SPTREMBL:O77815"
/translation="MGETLGDSLIDPESDAFDLTSANI SOEVTWDTMEPWPNGFI SSVLSVMDHSHFEDIKPPTVDFSSISTPHYEDIPFRADPVAADYKVDLKLDYQ SAIVLEVSPYSEKIQLYNKPHEEPSNSLMAIECRVGDKASGFHYGVACGKG FFRRTIRLKLIDRCDLNCRIHKSRNKCOCYCRQKCLAVGSHNAIRFGMPAEKE KLAIEISDIDLPESADLRALKHLDYSIKRSPILAKARAILITGKTDKSPVY YDNMSLMGEBKIKFKHTTPQEOSKEVAIRFGCCPRSENAOETTEAKNIPGY NLDNDVTLTKYVHEIITTMASLMKMDGVLISBEOGFEWTEFLSLRKPDEME KFEFAVKFNALDDSDLAFLAVIILSGRPRLNVKPIEDIQDMLLQALEYQLK NHPSSQLFALKQMTDLRQIVTEHVQLQVIKETEDMSLHPLQEIYKDLV"

BASE COUNT
459 a 408 c 382 g 391 t

ORIGIN

Alignment Scores:
Pred. No.: 3.03e-154 Length: 1640
Score: 2160.50 Matches: 430
Percent Similarity: 86.53% Conservative: 7
Best Local Similarity: 85.15% Mismatches: 5
Query Match: 61.59% Indels: 63
DB: Gaps: 2

US-09-931-007a-1 (1-688) x SSC6757 (1-1640)

Oy 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20
|||||
Db 61 ATGGGTGAACCTCTGGGAGATTCTCTTATTGACCCAGAAAGGATGCTTC--GACACG 117

Oy 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrPro 40
|||||
Db 118 CTGCTGCAACATTTCACAGAGGTGACCATGCTTGACACCGAATGCCGTTTGCGCC 177

Oy 41 ThrAsnProGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
|||||
Db 178 ACCAATTGGATCAGCTGCTGACCTGCGGATGAGGACGACCATCCACCTCTT 237

Oy 61 AspIleIleSerPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80
|||||
Db 238 GACATCAAGCCCTTACACAGCTGTTGATTCACGATTCACCTCCACACTATGAAGAC 297

Oy 81 IleProPheThrArgThrAspProValAlaAspTyrIleTyrAspLeuIleGluGln 100
|||||
Db 298 ATCCCATTCGCGAGAGCTGATCCAAATGTTGAGATTAAGTAGACCTGAAGCTCCAG 357

Oy 101 GluTyrGlnSerAlaIleIleValGluProAlaSerProProTyrTyrSerGluIleThr 120
|||||

Db 358 GACTACCAAGATGCCATCAAAAGTGAGGCGCTGTGTCCCGCTTATTATTCTGAAGAACT 417
Oy 121 GluLeuTyrAsn----- 124
|||||
Db 418 CAGCTGTACAAACACCTCAGCAAGACCTTCCAACTCCCTCATGGCAATTGATGCCGT 477
Oy 124 ----- 124
Db 478 GTCTGGGGAATAAAGCCGCGGCTTCCACTATGAGATTGATGTTGAAGATGCAAG 537
Oy 124 ----- 124
Db 538 GATTTCCTCCGAGGACATCATGATTGAGCTTATTATGATAGGTGGCATTTAACTGT 597
Oy 125 -----ArgAsnIleCysGlnTyrCysArgPheGlnIleCysLeuAla 138
|||||
Db 598 CGGATCCCAAAAAAGTAAATTAATGATGATCTCGTTCGTTTCAAAATGCTCTCT 657
Oy 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluIleGluIle 158
|||||
Db 658 GTGGGATGTCTCATTAAGCCATCATGAGTTTGGCGGATGCCACAGAGCTGAGAGAGAG 717
Oy 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
|||||
Db 718 CTGTTGGCAGATCTCCAGTATGATGATGACACGAGTGAACCCAGAGATGCTGACCTCCG 777
Oy 179 AlaLeuAlaIleLeuTyrAspSerTyrIleIleCysSerPheProLeuThrIleValAla 198
|||||
Db 778 GCGCTGGCAAGACCTTGTATGACTCATATCAATAAGCTCTTCCCTACCAAAAGAC 837
Oy 199 AlaArgAlaIleLeuThrGlyIleTyrThrIleAspIleSerProPheValIleTyrAspMet 218
|||||
Db 898 GCGAGGGGATCTTACAGCAAGAAAGCAGACAAATACACCTTTGTATCATATACATG 897
Oy 219 AsnSerLeuMetMetGlyGluAspIleIleCysPheIleIleTyrPheProLeuGlnIle 238
|||||
Db 898 AATTCCTTAAGATGGAGAGATTAATCAAGTTCAAAACATCACCCCTCGAGAGAG 957
Oy 239 GlnSerIleGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258
|||||
Db 958 CAGACCAAGAGGTGGCCATTCGCAATCTTTCAGGGGTGTGATTCGCTCGTGGAGGCC 1017
Oy 259 ValGlnGluIleThrGluTyrAlaIleCysSerIleProGlyIlePheValAsnLeuAspLeuAsn 278
|||||
Db 1018 GTGCAGAGATCACAGAGATATGCCAAGAACATCCCTGGGTTTGTAAACCTTGACCTGAT 1077
Oy 279 AspGlnValIleThrLeuLeuIleTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
|||||
Db 1078 GACCAAGTAACTCTCTAAAGTATGGCGTCCACGAGATCATTTACCATGCTGCTGCC 1137
Oy 299 LeuMetAsnIleAspGlyValIleuIleSerGlnGlyGlnGlyPheMetThrArgGluPhe 318
|||||
Db 1138 TTGATGATTAAGACGGGGGTCTCTATCTCCAGGGCCCAAGGATTCATATCAAGAGGATT 1197
Oy 319 LeuIleSerLeuArgIleProPheGlyAspPheMetGluProIleCysPheGluPheAlaVal 338
|||||
Db 1198 CTAAGAGACCTCAGAAACCCCTTGGTGTACTTATGAGGCCCAAGTTCGATTTGCTGTG 1257
Oy 339 LysPheAsnAlaLeuGluLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle 358
|||||
Db 1258 AAGTTCAAAGCAGCTGCAATTTGATGACAGCACTGGCGGATATTATAGCTGTCATATT 1317
Oy 359 LeuSerGlyAspArgProGlyLeuLeuAsnValIleLysProIleGluAspIleGluAspAsn 378
|||||
Db 1318 CTCAGTGGAGACCGCCGCTTGTGTAATGTGAAGCCCATGAGAGCAATTCAGACATAT 1377
Oy 379 LeuLeuGlnAlaLeuGluLeuGlnLeuIleCysLeuAsnHisProGluSerSerGlnLeuPhe 398
|||||
Db 1378 TTGCTGCAAGCCTTGGAGCTGACGTCAAGTTGAACCAACCCAGAGTCTCCACACTCTTT 1437
Oy 399 AlaIleLeuLeuGlnIleCysMetThrAspLeuArgGlnIleValIleThrGluHisValGlnLeu 418
|||||
Db 1438 GCTAAGCTGTCTCCAGAAATATGACAGACCTCAGACAGATTGTGACAGAGACGTGACGCTG 1497

QY 419 LeuGlnValIleIleLysLysThrGluThrAspMetSerLeuHisProLeuGlnGluIle 438
|||||
Db 1498 CTGCAAGTATATAAGACAGACAGACGACATGACCTTACCACCTCCTACAGGAATA 1557

QY 439 TyrLysAspLeuTyr 443
|||||
Db 1558 TACAAGGACTGTAC 1572

RESULT 12
BC021798 1782 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, peroxisome proliferator activated receptor gamma,
DEFINITION clone MGC:18439 IMAGE:4191539, mRNA, complete cds.
ACCESSION BC021798
VERSION BC021798.1 GI:18255315
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1782)
REFERENCE Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
REMARK Contact: MGC help desk
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amgdbcm.tmc.edu
Guanarane, P.H., Garcia, A.M., Lu, X., HuJYK, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAX Plate: 24 Row: i Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6755137.
Location/Qualifiers
1. 1782
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="PVB/N"
/clone="MGC:18439 IMAGE:4191539"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_id="NCI CGAP_SG2"
/lab_host="DH10B"
/note="vector: pCMV-SPORT6"
46. 1563
/codon_start=1
/product="peroxisome proliferator activated receptor
gamma"
/protein_id="AAH21798.1"
/db_xref="GI:18255316"
/db_xref="locusID:19016"
/translation="MGETLGDSVPDEHGAFADALPMSTSGEITWVDTMPFWPTNFG
ISSVLSMEDSHSFDIKPTTVDFSSISAPHEDIPTFADPMVADKYKDLKQEX
QSAIKVEPASPYPYSEKTLNRPHEESNLMAIECVCDKASGFVGYHAGEGCK
GEFRTIRKLIDRCDLNCRHKSRKRCYCRKCLAVGMSNATRFGRMOAOK
EKLAETSSDIDLPESADRLAKLHYDSYISFPTTKAKAAILTKRTDMSPV
IYDMSLIMGEDKIKFKHITPLOSKEVALIRIQCOCFRSEVAOETTERAKINPGF
INLNDQVTLTKYGVHEITVTMLASLNNKQGVLSISQGRMTREFLKSLRPGDFM
EPKEFAVKNFALELDSDLAIFLAVIILSGDRQLNVRKPIEDIDQNLQALEIQLK

FEATURES
source

CDS

LNHPESOLFAPKVLQKMTDLROIIVTEHVOLLHVIKKTEPDMSLPILQEIYKDY"
BASE COUNT 533 a 416 c 386 g 447 t
ORIGIN
Alignment Scores:
Pred. No.: 7,33e-154 Length: 1782
Score: 2156.00 Matches: 425
Percent Similarity: 85.74% Conservative: 8
Best Local Similarity: 84.16% Mismatches: 10
Query Match: 61.46% Indels: 62
DB: 10 Gaps: 1
US-09-931-007a-1 (1-688) x BC021798 (1-1782)
QY 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerThrAspThr 20
|||||
Db 46 ATGGGGAACCTGGGAATTCCTCTGTGACCAAGACATGTCCTTCCTATGCA 105
21 LeuSerAlaAsnIleSerGlnIleuThrMetValAspThrGluMetProThrPro 40
|||
Db 106 CTGCCTATGACGACCTTACACAAATTTACATGTTGACACAGATGCTATGCGCC 165
41 ThrAspPheGlyTlSerSerValAspLeuSerValMetGluAspHisSerHisPhe 60
|||||
Db 166 ACCAATTCGGAATCAGCTCTGACACCTCTCCGATGGAACACACGCAATTCCTTT 225
QY 61 AspIleLysProPheThrValAspPheSerSerIleSerThrProHisThrGluAsp 80
|||||
Db 226 GACATCAAGCCCTTTCACACAGCTGATTTCTCAGCATTTGCTCCACACTATGAAGAC 285
QY 81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGln 100
|||||
Db 286 ATTCATTTACAAAGAGCTACCCATGTTGCTGATTAACAAATATGACCTGAAGCTCCA 345
QY 101 GluTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluYthr 120
|||||
Db 346 GAATACCAAGTCGATCAAGATAGAACCTGCATCTCCACTTATATTCGAAAAAGAC 405
QY 121 GlnLeuTyrAsp 124
|||||
Db 406 CAGCTCTACAAAGAGCCCTCATGAAGAACCTTCACTCCCTCATGAGCCATTGATGCCGA 465
QY 124 124
Db 466 GTCTGTGGGGAATAAGCATCAGGCTTCCACTATGAGATTCATGCTTGTGAAGATGCAAG 525
QY 124 124
Db 526 GGTTCCTTCCGAAGACCATCCGATTGAAGCTTATTTATGATAGGTGATCTTAACCTGC 585
QY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
Db 586 CGGATTCACAAAAAGATGAAATGAATGTCAGTCTGCTGGTTTCACAAAGTCCCTCTCT 645
QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158
|||||
Db 646 GTGGGATGTCACAAATGCCATCAGGTTTGGCGGATGCCACAGCCGAGAGAGAGAAAG 705
QY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
|||||
Db 706 CTGTTGGCGGAGATCTCCAGTATATTCGACCGAGAACCCAGAGTCTGATCTGCGGA 765
QY 179 AlaLeuAlaLysIleLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198
|||||
Db 766 GCCCTGGCAAGCATTTGTATGACTCATACATAAAGTCTTCCCTGACCAAGCCAAAG 825
QY 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218
|||||
Db 826 GCGAAGGCGATTTTACAGAGGAAGAACACGCAATATCACCATTGTCTATCAGACATG 885
QY 219 AsnSerLeuMetMetCylGluAspLysIleLysPheLysHisIleThrProLeuGlnGlu 238
|||||
Db 886 AATTCCTTAATGATGGAGAGATAAATCAAGTTCAAAACATATACACCCCTCGAGAG 945

OY	239	GlnSerLysGluValAlaAlaIleArgIleIlePheGlnGlyCysGlnPheArgSerValGluAla	258
Db	946	CAGACCAAAAGAGGTGGCCATCCGATATTTTTCAAAGGTTGCCACTTTCGATCCCTAGAACGC	1005
OY	259	ValGlnGluIleIlePheGlnIleTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn	278
Db	1006	GGCGAAGAGATCCACAGAGATGCCCCAAAATATCCCTGGTTTCATTAAACCTTGATTTGAAAT	1065
OY	279	AspGlnValIleThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer	298
Db	1066	GACCAAGTACCTCTCTCAAGATATGTGTCCATGAGATCATTTACACGATGCTGGCCTCC	1125
OY	299	LeuMetAsnLysAspGlyValIleuIleIleSerGluGlyGlnGlyPheMetThrArgGluPhe	318
Db	1126	CTGATGAATATAAGATGAGAGTCTCTCATCTCCAGAGGCCCAAGATTCATGACCAGGAGATTCC	1185
OY	319	LeuLysSerIleuArgLysProPheGlnLysPheMetGluProLysPheGluPheAlaVal	338
Db	1186	CTCAAAACCTCTGGAGAACCCCTTTGGTGACTTATGAGACCTTAAGTTTGAGTTTCTGTGG	1245
OY	339	LysPheAsnAlaIleuGluLeuAspSerSerAspLeuAlaIlePheIleAlaValIleIle	358
Db	1246	AAGTTCATTCGACCTCGGATTTAGATACAGTAGACTTGCGTATTTATATAGCTCTCATTTATT	1305
OY	359	LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn	378
Db	1306	CTCACTGGAGAGACCGCCAGGCTTCTGTGAACGGAGGCCCATCGAGACATCCAMAAGAACAC	1365
OY	379	LeuLeuGlnAlaIleuGluLeuGlnIleuLysLysLeuAsnHisProGlnSerSerGlnLeuPhe	398
Db	1366	CTGCTGCAGAGCCCTTGGAACCTGCAGCTCAACCTGGAATTCACCCAGAAATCCTCTCAGCTGTTC	1425
OY	399	AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValIleThrGluHisValGlnLeu	418
Db	1426	GCCAGAGTGCTCCAGAAATGACAGACCTCAGCGCAGATGCTCACAGAGACGTCGACGCTA	1485
OY	419	LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnIleIle	438
Db	1486	CTGCATGTGATCCAAAGAACACAGACAGACATGAGCTTCACCCCTCTCCAGGAGATGC	1545
OY	439	TyrlLysAspLeuTyr 443	
Db	1546	TTCACAGGACTTGTAT 1560	
RESULT 13			
LOCUS	MMU09138	1780 bp	linear ROD 18-FEB-1995
DEFINITION	Mus musculus peroxisome proliferator activated protein-gamma-2		
ACCESSION	U09138		
VERSION	U09138.1	GI:500639	
KEYWORDS	.		
SOURCE	Mus musculus.		
ORGANISM	Mus musculus.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 (bases 1 to 1780)		
TITLE	Tontonoz,P., Hu,E., Graves,R.A., Budavari,A.I. and Spiegelman,B.M.		
JOURNAL	mPAR gamma 2: tissue-specific regulator of an adipocyte enhancer		
MEDLINE	Genes Dev. 8 (10), 1224-1234 (1994)		
PMID	95011536		
REFERENCE	7926726		
AUTHORS	2 (bases 1 to 1780)		
TITLE	Hu,E.		
JOURNAL	Direct Submission		
PMID	Submitted (21-APR-1994)		
REFERENCE	Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,		
FEATURES	Location/Qualifiers		
Source	1..1780		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		

CDS

```

/cell_line="3T3-F442A"
/tissue_type="adipose tissue"
/clone_id="lambda ZapII library"
40..1557
/codon_start=1
/product="peroxisome proliferator activated
protein-gamma-2"
/protein_id="AA062277.1"
/db_xref="gi:500640"
/translation="MGETLGDSPVDPFHGAFAADALPMSTQELTMVDTMPFNPFTNG
ISSVDLSVMEDHSFDIKPTPTVDSESIAPHEVDLPTFRADRPADYKYDLKLOEY
OSALIKVEPASPSSEKTOLYNRPHEPSSNSIMLIERGVGRASGPHVCVACECK
GFFRRTRFKLITRDCLNCRHKSRNKOYCRFOCLAVGASHNAIRGRMPQARK
EKLAEIISDDQLNPESADIRALAKLHYDSYIKSPYLAKARALITGTTKDSPV
IYDNNSLMMGEDKIKFHIITPLOSKEVALIRFOGOFSEVAVOETETAKNIPOF
INLDNDQVTLKYGVEHIIYTMASLMMNKGVILISGOGFMPEEFLKNLRPGEDM
EPKEPFAKFNALFLDSDSLAIFLAVAILITSDRGILNVKPIEDIONLLOALELOLK
LNHESSOLFPAKVIQKMTDLRQIYTHENVOLLHVIKRTETMSLHPLQEIYKDLX"

BASE COUNT      531 a      412 c      387 g      450 t
ORIGIN

Alignment Scores:
Pred. No.:      1,23e-153      Length:      1780
Score:          2153.00      Matches:      424
Percent Similarity: 85.74%      Conservative: 9
Best Local Similarity: 83.96%      Mismatches: 10
Query Match:      61.37%      Indels:      62
DB:              Gaps:      1

US-09-931-007A-1 (1-688) x MMU09138 (1-1780)

QY      1 MetGlygluThrLeuGlyAspSerProIleasProGluSerAspSerPheThrAspThr 20
|||||
40 ATGGGTGAAGACTCTGGAGATCTCTCGTGTGACCCAGACATGGCGCTTCGTGATCA 99
|||||

QY      21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrPro 40
|||||
100 CAGCCTATGAGCATTCCACAAAGAAATTCACAGTGTGACACAGACATCCATTCTGGGCC 159
|||||

QY      41 ThrAspNheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
|||||
160 ACCAATTCTGGGATAGCTCTGTGGACCTCTCCGTGATGGAAGACCACTCCGATTCCTTT 219
|||||

QY      61 AspIleLeuProPheThrThrValAspPheSerIleSerThrProHisTyrGlnAsp 80
|||||
220 GACATCAAGCCCTTACACAGATGATTCTTCCACACATTCTGCGTCCACACTATGAAAGAC 279
|||||

QY      81 IleProPheThrAspThrAspProValAlaAlaAspTyrIleTyrAspLeuIleGln 100
|||||
280 ATTTCATTCCACAGAGCGTCGACCAATCGTTCGTCAATTAATGACCTCAAGCTCCAA 339
|||||

QY      101 GluTyrGlnSerAlaIleLeuValGluProAlaSerProProTyrTyrSerGlnIleThr 120
|||||
340 GAATACCAAGAGTGCATCAAGTAAAGTAAACCTGCATCTCCACCTTATTATTGAAAAGACC 399
|||||

QY      121 GlnLeuTyrAsn----- 124
|||||

DB      400 CAGCTCTCAACAAGCGCTCATGAAGAACCTTCTACTCCCTCATGGCCATCGAGAGCCGA 459
|||||

QY      124 ----- 124

DB      460 GTCTGTGGGATAAAGCATCAAGCTTCCACTATATGACATTATGCTTGTGAAGATGCAG 519
|||||

QY      124 ----- 124

DB      520 GGTATTTTCCAGACACATCCGATTGAAGCTTATTATATAGTATAGGTGTGATTAACCTCC 579
|||||

QY      125 -----ArgAsnIleScysGlnTyrCysArgPheGlnIleScysLeuAla 138
|||||

DB      580 CGGATCCACAAGAAAGTAAATTAATAGTCAGTAACTGTGGTTTTCAGAACTGCGCTTGCT 639
|||||

QY      139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluIleGluIle 158
|||||

```

D	640	GTGGGAGTCTCACAATGCCATCAGGTTTGGCGGATGCCAGGCCAGAGAGANG	699
O	159	LeuLeuAlaGluIleSerSerAspIleAspGluLeuAsnProGluSerAlaAspLeuArg	178
D	700	CTGTTGGCGGAATCTCCAGTATATCGACCGAGCTGAACCCAGAGTCTCTATCTGGCA	759
O	179	AlaLeuAlaIleHisLeuThrAspSerTyrIleLeuSerPheProLeuThrLysAlaLys	198
D	760	GCCCTGGCAACGATTTGTATGATCATATCAATAAAGCTCTCCCGGTGACCAAGCAG	819
O	199	AlaAlaGAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet	218
D	820	GCGAGGGCATGTTTACAGAGAAAGACAAAGGAGCAATCCATTGTCTATCGACAG	879
O	219	AsnSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGlu	238
D	880	AATTCCTTATATGATGGAGAAATATAATCAAGTTCAAACATATCACCCCTCGCAGAG	939
O	239	GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla	258
D	940	CAGACCAAAAGAGTGGCCATCCGAATTTTCAAGGGTGGCAGTTTCGATCGTAGAGCC	999
O	259	ValGlnGluIleThrGluThrLysSerIleProGlyPheValAsnLeuAspLeuAsn	278
D	1000	GTGCAGAGATCACAGAGATATGCCAAAATATCCCTGGTTTCATTAAACCTGATTTGAAT	1059
O	279	AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer	298
D	1060	GACCAAGTACTCTGCTCAGATATGCTGTCATGAGATCATCTACACGATGCTGGCTCC	1119
O	299	LeuMetAsnLysAspGlyValLeuLeuSerGlnGlnGlnGlyPheMetThrArgLuphe	318
D	1120	CTGATGAATAAAGATGAGATGCTCATCTCAGAGGGCCAGATTCATGACCAAGGGATT	1179
O	319	LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal	338
D	1180	CTTCAAAAACCTCCGGAGGCCCTTGTGCTTATGAGACCTTAAGTTTGAAGTTTGGCTGT	1239
O	339	LysPheAsnAlaLeuGluLeuAspSerAspSerAspLeuAlaIlePheIleAlaValIleIle	358
D	1240	AGTTTCAATGCACTGGAATTAATGACAGTGAAGTGGCTGATTTTATGCTGCTATATT	1299
O	359	LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn	378
D	1300	CTCATGTGAGACGCCGACAGGCTTGCTGAAGCTGAAGCCATCGAGAGATCCAAAGACAC	1359
O	379	LeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPhe	398
D	1360	CTGCTGCAGGCCCTGGAAGCTGCAAGCTCAAGCTGAATCACCCAGAGTCCCTCAGCTGTT	1419
O	399	AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeu	418
D	1420	GCCAGGTGCTCCAGAGATGACAGACCTCAGCAGATGTGCACAGAGACGTCGACGTA	1479
O	419	LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIle	438
D	1480	CTGCATGTGATCAAGAGACAGAGACAGATGAGAGCTTACACCCCTGCTCCAGAGATC	1539
O	439	TyrLysAspLeuTyr 443	
D	1540	TTCAAAGGACTGTAT 1554	
RESULT 14			
LOCUS	AB019561	1518 bp	mRNA linear ROD 08-JAN-1999
DEFINITION	Rattus norvegicus mRNA for PPAR gamma2, complete cds.		
ACCESSION	AB019561		
VERSION	AB019561.1	GI:4115708	
KEYWORDS	PPAR gamma2.		
SOURCE	Rattus norvegicus (Strain: Sprague-Dawley) 8-week-old male adipose tissue cDNA to mRNA.		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
AUTHORS	1. (sites)	
TITLE	Tanaka,T. and Itoh,H.	
JOURNAL	Down-regulation of PPAR gamma	
REFERENCE	Unpublished	
AUTHORS	2 (bases 1 to 1518)	
TITLE	Itoh,H., Tanaka,T. and Nakao,K.	
JOURNAL	Submitted	
FEATURES	Submitted (12-AUG-1998) Hiroshi Itoh, Kyoto University Graduate School of Medicine, Dept of Med and Clin Sci; 54 Shogoin Kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan (E-mail: hito@kuhp.kyoto-u.ac.jp, Tel:81-75-751-3170, Fax:81-75-771-9452)	
source	Location/Qualifiers	
	1..1518	
	/organism="Rattus norvegicus"	
	/strain="Sprague-Dawley"	
	/db_xref="taxon:10116"	
	/sex="male"	
	/tissue_type="adipose tissue"	
	/dev_stage="8-week-old"	
CDS	1..1518	
	/codon_start=1	
	/product="PPAR gamma2"	
	/protein_id="BAA36485.1"	
	/db_xref="GI:4115709"	
	/translation="MGETGDPDPDEPGHAFADALPMSTSEITMWTMPFRTNG ISSVDLSMDHSHSFDIKFTTYVDSSISAPHEIDIEPTRADPVADYKDLQLEY QSAIKVEPASPYYSEKTOLYNRHESRNSLMAIECRVCGDKASGHYGVHACEGK GFPRRTILRLIYRCDLNCRIHKSRNKCQYCFPOCLAVGSHNMAIRFGMRQAEK EKLAEISSDIDOLNPEASDLRALAKHLUDYIKSFPLTKAKARAILGKTTDKSPFY IYDNNLSMGEDIKFKHPIIPLOESKEVATIRFOGQCFSEVAVOITTEYAKRIPGF INLDNVOYTLKRYGHEITTYTMAISLNNKGVILSBOGCTNTEFLKSLKPRPGDM EPKFEFAVFNALDELDDSLATFIYIILSDRGLNVRPEDIQNLQALELQK LNHPESSOLFPAKVLOKMTDLROIYETEHVOLHVIKTKETDMSLHPLOEYKDLY"	
BASE COUNT	422 a 379 c 358 g 359 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	2.42e-153	
Score:	2148.00	
Percent Similarity:	85.54%	
Best Local Similarity:	83.76%	
Query Match:	61.23%	
DB:	10 Gaps: 1	
US-09-931-007A-1 (1-688) x AB019561 (1-1518)		
OY	1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr	20
D	1 ATGGGTAACACTCGGAGATCTCTGTTGACCCAGAGCATGAGTCCCTCCATGCA	60
OY	21 LeuSerAlaAsnIleSerGlnLysMetValAspThrGluMetCProPheTrPro	40
D	61 CTGCCTATGAGCAGCTTCAAGAAATACCATGTTGACACAGAGATGCCATTCGGCCC	120
OY	41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisPhe	60
D	121 ACCAAGCTTGGAAATCAGCTGTGGACCTCTGTGATGAGATGACACCTCCATTCCTT	180
OY	61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp	80
D	181 GACATCAAAACCTTTACACGAGTGTATTCACAGATTTCTGCTCCACACATATGAAGAC	240
OY	81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGln	100
D	241 ATCCCGTTACAAAGAGTGAACCAATGCTGTGATTCAATAATATGACCTGAAGCTCAA	300
OY	101 GltTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGlyLysThr	120
D	301 GAATACCAAGTGCATCAAGTACAGCTCGCTCCGCCCTATTATTTCGAAAAAAC	360

```

Qy 121 GlnLeuTyrAsn-----124
    |||||
Db 361 CAACTCTACAAAGGCGCACATGAAGAGCCTTCAACCTCCATGCGCATGAGAGTCCGA 420
Qy 124 -----124
Db 421 GTCTGTGGGGATAAAGCATCAGGCTTCCACTATGAGAGTCATGCTTGTGAAGATGCAAG 480
Qy 124 -----124
Db 481 GGTTTTCCGAGAACCATCCGATTTGAGCTTATTTGATAGTGATCTTAACTGT 540
Qy 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
    |||||
Db 541 CGGATCCACAAAAGAGATAGAAATTAATGTGACGTCTGCGTTCCAGAAAGTCCCTGCT 600
Qy 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158
    |||||
Db 601 GTGGGGATGTCTCACAAAGCCATCAGGTTTGGCGCAATGCCACAGCCGAGAGAGAGAG 660
Qy 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
    |||||
Db 661 CTGTTGGCGGAGATCTCCAGTGATATTCACAGCTGAACCCAGAGTCTGCTGATCTGCCA 720
Qy 179 AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198
    |||||
Db 721 GCCCTGGCAAGCATTTGTATGACTCATACATAAAGTCTCCCGCTGACCAAGCCCAAG 780
Qy 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218
    |||||
Db 781 GCGAGGGGATCTTGCACAGGAAGCAACACAAATACCATTTGTCATCTACACATG 840
Qy 219 AsnSerLeuMetMetGlnGluAspLysIleLysPheLysHisIleThrProLeuGlnGlu 238
    |||||
Db 841 AATTCCTTAATGATGGGAGAGAACAAATCAAGTTCAACATATACCCCTCGCAGAG 900
Qy 239 GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258
    |||||
Db 901 CAGACCAAAAGAGTGGCCATCCGCTTTTCAAGGCTCCAGTTTCGATCCGTGGAAGCT 960
Qy 259 ValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278
    |||||
Db 961 GTGCAAGAGATCAGAGATATGCCAAAATATCCCTGTTCACTTAACCTTGACTTGAAAT 1020
Qy 279 AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
    |||||
Db 1021 GACCAAGAGACTCTGCTCAAGTATGCTTCCATGAGATCATCTACACCATGCTGCTTCC 1080
Qy 299 LeuMetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGluPhe 318
    |||||
Db 1081 CTGATGATTAAGATGAGTCTCTCATATCAGAGGACCAAGATTCATGACCAAGGAGTTC 1140
Qy 319 LeuLysSerLeuAlaGlyLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338
    |||||
Db 1141 CTCAAAACCCGCGGAAACCCCTTGGTACTTATGAGAGCCCTAAGTTTGAGTTTGCCTGTG 1200
Qy 339 LysPheAsnAlaLeuGlnLeuAspSerSerAspLeuAlaIlePheIleAlaValIleIle 358
    |||||
Db 1201 AAGTTTCATGACACGTGATTTGATGACAGTACCTTGCCCATATTTATAGCTGTCTTATT 1260
Qy 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
    |||||
Db 1261 CTCAGTGGAGACCGCCGCGCTGCTGTAAGCTGAAGCCCATCGAGAGATCCAGACAAAC 1320
Qy 379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluLysSerGlnLeuPhe 398
    |||||
Db 1321 CTGCTGCGAGGCCCTGGAATCTCAGCTGAAGCTGAACCCCGGAGTCTCCACAGCTGCTTC 1380
Qy 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValIleThrGlnHisValGlnLeu 418
    |||||
Db 1381 GCCAAGAGTGTCTCCAGAGATACAGACCTCAGGACGATTTGTCACAGAGACAGTGCAGCTA 1440
Qy 419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGlnIle 438

```

```

Db 1441 CTGCTGATCAAGAGAGACGAGACAGATATGAGACCTTCAACCTCTGCTCAGAGATC 1500
Qy 439 TyrLysAspLeuTyr 443
    |||||
Db 1501 TACAAGGACTTGTA 1515

RESULT 15
LOCUS AF156666
DEFINITION Rattus norvegicus peroxisome proliferator-activated receptor gamma
ACCESSION AF156666
VERSION 2 (ppargamma2) mRNA, complete cds.
KEYWORDS AF156666.1 GI:5107414
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1785)
AUTHORS Guardiola-Diaz,H.M., Rehmark,S., Usuda,N., Albrektzen,T.,
Feltkamp,D., Gustafsson,J.A. and Alexson,S.E.
TITLE Rat peroxisome proliferator-activated receptors and brown adipose
JOURNAL tissue function during cold acclimatization
MEDLINE 99367468
PUBMED 10438514
REFERENCE 2 (bases 1 to 1785)
AUTHORS Guardiola-Diaz,H.M., Rehmark,S., Usuda,N., Albrektzen,T.,
Feltkamp,D., Gustafsson,J.A. and Alexson,S.E.H.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-1999) Division of Clinical Chemistry, Karolinska
INSTITUTE, C1-74 Huddinge Hospital, Huddinge S-141 86, Sweden
FEATURES
source
1..1785
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1..1785
/gene="ppargamma2"
49..1566
/gene="ppargamma2"
/note="nuclear receptor"
/codon_start=1
/product="peroxisome proliferator-activated receptor gamma
2"
/protein_id="AAD40119.1"
/db_xref="GI:5107415"
/translation="MGETLGDPPVDPEHGAADALPMSTSOEITMVDTEMPFPTNG
ISSVDLSVMDHSHSPDIKPTVDYDFESIASPHYEDIPTRADPMVADYKVDLKLQFY
OSAKYVPASPPYSEKTOLYNRHPEPSNMAIERVCGDRASGTHYGVHACEGK
GFERRTIRKLITDRCDLNCRIHKSNNKQCYCFQCLAVGSHNAIRGMPQARK
EKLALAISSDIDOLNPESADLRALAKLYDSYIKSFPLTAKARALITGTTDKSPV
IYDMSNLMMEDKIKFKHITPLOSKEVALIRIQQGCFSEVAVQETFEYAKNIPGF
INDLDQVTLTKGVHEIITYMLASLMMKGVILISGSGFMREFFLSRKPGDFM
EPKPEFAKFNALIEDDSDLAIFLAVITILISGDRGLILNVKPIEDIONLLOALELOK
LNHPESQALFKAKYOKMTDLROTYTEHVQLLHVITKRTDMSLHPLQLEIYKDIY"
BASE COUNT 533 a 417 c 395 g 440 t
ORIGIN
Alignment Scores:
Pred. No.: 2,95e-153 Length: 1785
Score: 2148.00 Matches: 423
Percent Similarity: 85.54% Conserved: 9
Best Local Similarity: 83.76% Mismatches: 11
Query Match: 61.23% Indels: 62
DB: 10 Gaps: 1
US-09-931-007a-1 (1-688) x AF156666 (1-1785)
Qy 1 MetGlyGlnThrLeuGlyLysSerProIleAspProGluSerAspSerPheThrAspThr 20
    |||||

```

```

Db 49 ATGGTGAACCTCTGGAGATCCCTCTGTGACCAGACATGTCCTCGTGATGCA 108
QY 21 LeuSerAlaAsnIleSerGlnIleMetThrMetValAspThrGlnMetProPhePro 40
Db 109 CTGCGCTATGACACTTCACAGAAATACCATGGTTGACACAGAGATGCACTTG6CCC 168
QY 41 ThrAsnPhgGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
Db 169 ACCAACTTGGATGATGACGCTGTGGACCTCTCGTATGGATGACACACCCCATTCCTT 228
QY 61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80
Db 229 GACATCAAAACCCCTTACCCAGGTGATTTCTCCAGCATTTCTGCTCACACTATGAAGAC 288
QY 81 IleProPheThrArgThrAspProValAlaAspTyrTyrAspLeuLysLeuGln 100
Db 289 ATCCCGTTACAAAGAGCTACCCCAATGTTGCTGATTAATAATATGACCTGAAGCTCCAA 348
QY 101 GlnTyrGlnSerAlaIleLysValGluProAlaSerProTyrTyrSerGluLysThr 120
Db 349 GAATACCAAAAGTGCATCAAAAGTAGAGCTGCGTCCCGCTTATTTCTGAAAAAAC 408
QY 121 GlnLeuTyrAsn----- 124
Db 409 CAAGCTTACAAAGGCGCACATGAAGACCTTCAAACTCCCTCATGGCATCGAGTCCGA 468
QY 124 ----- 124
Db 469 GTCGTGGGGATAAAGCATCAGGCTTCCACTATGGAGTGCATGCTTGTAAGATGCAAG 528
QY 124 ----- 124
Db 529 GGTTTTTCGGAAGAACCATCCGATGAGCTTATTATGATAGTGTGATCTTAACGT 588
QY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
Db 589 CGGATCCACAAAAAGATAGAAATGAATATGTCAGTACTGCGGTTTCAGAAAGTCCCTGCT 648
QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaLysGluLys 158
Db 649 GTGGGGATGCTCTCACATGCCATCAGTGTGGGCGCATGCCACAGGCCAGAGAGGAAG 708
QY 159 LeuLeuAlaGlnIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
Db 709 CTGTTGGCGGAGATCTCCAGTATTCGACAGCTGAACCCAGAGCTGCTGATCGCGA 768
QY 179 AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198
Db 769 GCCCTGGCAAGCATTTGTATGACTCATATAAGTCCCTCCGCTGACCAAGCCCAAG 828
QY 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218
Db 829 GCGAGGGCATCTTGACAGGAAGACACAGACAAATCACCATTGTCTATACGACATG 888
QY 219 AsnSerLeuMetMetGlyLysAspLysIleLysPheLysHisIleThrProLeuGlnGlu 238
Db 889 AATTCCTTAATATGAGGAGAAAGACAAATCAAGTTCAAACATATACACCCCTGCAGAG 948
QY 239 GlnSerLysGlnValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258
Db 949 CAGACCAAAAGAGGTGGCCATTCGCAATTTTCAAGGTGCGCAGTTTCGATCCGTGAAGCT 1008
QY 259 ValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278
Db 1009 GTGCAAGAGATACAGAGATATGCCAAAATATCCCTGGTTTCAITRACTTGACTTGAAT 1068
QY 279 AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
Db 1069 GACCAAGTACACTCTGCTCAAGATATGTTGCCATGAGATCATCTACACCATGCTGGCTCC 1128
QY 299 LeuMetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGluPhe 318
Db 1129 CTGATGAATAAAGATGAGTCTCTCATATCGAGGAGCAAGGATTCATGACAGGAGGATTTC 1188

```

```

QY 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338
Db 1189 CTCAAAAGCTCTGGAGAGCCCTTTGGTGCATTTATGGAGCTTAAGTTGAGTTGCTGTG 1248
QY 339 LysPheAsnAlaLeuGlnLeuAspSerAspLeuAlaIlePheIleAlaValIleIle 358
Db 1249 AAGTTCAATGACACTGGAATTAGATGACAGTGAAGTGGCCATATTATAGCTGCATTAT 1308
QY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
Db 1309 CTCAGTGGAGACCGCCAGAGCTTGTGAACGTGAAGCCCATGAGACATCCAAAGACAA 1368
QY 379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPhe 398
Db 1369 CTGCTGCAGGCCCTGGAATCCAGCTGAAGCTGAACCAACCCGAGATCTCCACAGCTGTT 1428
QY 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeu 418
Db 1429 GCCAAGGTCTCCAGAGATGACAGACCTCAGCGAGATTGTCACAGAGCAGCTGCAGCTA 1488
QY 419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIle 438
Db 1489 CTGCAATGTGATCAAGAGAGAGAGACAGATATGAGCTTCACCCCTGCTCCAGAGATTC 1548
QY 439 TyrLysAspLeuTyr 443
Db 1549 TACAAGACTTGAT 1563

```

Search completed: February 25, 2003, 05:32:08
Job time : 4897 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2003, 03:23:46 ; Search time 75 Seconds
(without alignments)
1890.139 Million cell updates/sec

Title: US-09-931-007a-1
Perfect score: 3508
Sequence: 1 MGETLGSPIDPESDSFTD.....KTEFMSLHPLQEIYKDLV 688

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SPTRMBL_21:*
2: sp_archaea:*
3: sp_bacteria:*
4: sp_fungi:*
5: sp_human:*
6: sp_invertebrate:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteria:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2212	63.1	505	6	Q95KZ8
2	2089	59.5	477	4	Q96J12
3	2064	58.8	475	6	Q95J17
4	2045	58.3	475	11	Q99PC7
5	1920	54.7	475	13	Q91878
6	1916	54.6	472	13	Q8QCGO
7	1698	48.4	477	13	Q98SFE
8	1564	44.6	393	13	Q9PU74
9	1549	44.2	393	13	Q9PU26
10	1325	37.8	333	13	Q9PW01
11	1312.5	37.4	532	13	Q9W712
12	1140	32.5	543	13	Q9OWP6
13	1117.5	31.9	443	13	Q918W3
14	1105	31.5	219	6	Q95KZ4
15	1089.5	31.1	468	6	Q95N78
16	1085.5	30.9	440	11	Q99ND3

17	1084.5	30.9	441	6	Q9MZG4	Q9MZG4 sus scrofa
18	1082	30.8	216	6	Q8WMS4	Q8WMS4 canis famill
19	1080.5	30.8	440	11	Q9QW10	Q9QW10 rattus norv
20	1078.5	30.7	521	13	Q918F6	Q918F6 salmo salar
21	1076.5	30.7	368	13	Q90Z66	Q90Z66 brachydanio
22	1076.5	30.7	440	11	Q62879	Q62879 rattus norv
23	1076.5	30.7	441	6	Q9N2H4	Q9N2H4 oryctolagus
24	1070.5	30.5	337	13	Q8UX11	Q8UX11 oryctias lat
25	1070	30.5	468	13	Q8QSC1	Q8QSC1 anser anser
26	1068.5	30.5	468	13	Q918W4	Q918W4 gallus gall
27	1067.5	30.4	436	6	Q9N128	Q9N128 bos taurus
28	1065.5	30.4	470	6	Q9N129	Q9N129 bos taurus
29	1056	30.1	328	13	Q9PU76	Q9PU76 crocodylus
30	1035.5	29.5	329	13	Q90Z67	Q90Z67 brachydanio
31	1033	29.4	467	11	Q70527	Q70527 cavia porce
32	1028.5	29.3	468	6	Q9N135	Q9N135 sus scrofa
33	1021.5	29.1	337	13	Q90Z62	Q90Z62 salmo salar
34	1020.5	29.1	304	13	Q9PU75	Q9PU75 crocodylus
35	1000	28.5	252	6	Q46479	Q46479 oryctolagus
36	969	27.6	482	13	Q918F7	Q918F7 salmo salar
37	935	26.7	436	13	Q90Y02	Q90Y02 petromyzon
38	802.5	22.9	361	4	Q9BUD4	Q9BUD4 homo sapien
39	738	21.0	230	6	Q9GLC8	Q9GLC8 mustela vis
40	715	20.4	196	13	Q9W718	Q9W718 gallus gall
41	684.5	19.5	233	13	Q8UUX2	Q8UUX2 pimephales
42	571.5	16.3	174	6	Q46477	Q46477 oryctolagus
43	509.5	14.5	157	4	Q75780	Q75780 homo sapien
44	394.5	11.2	549	13	Q90970	Q90970 gallus gall
45	392.5	11.2	615	11	Q922A5	Q922A5 mus musculu

ALIGNMENTS

RESULT 1

Q95KZ8 PRELIMINARY; PRT; 505 AA.
ID Q95KZ8
AC Q95KZ8;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Peroxisome proliferator-activated receptor gamma 2.
GN PPARGAMA2.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhou J., Wilson K.M., Medh J.D.;
RT "Identification of novel peroxisome proliferator-activated receptor-
gamma splice variants and induction of PPAR-gamma expression by a
RT high-fat diet in monkey macrophages."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL: AY048695; AI05262.1; -
DR InterPro: IPR000536; Hormone_rec_11g.
DR InterPro: IPR001628; ZnF_C4steroid.
DR Pfam: PF00104; hormone_rec_1.
DR Pfam: PF00105; zf-C4; 1.
DR ProDom: PD000035; ZnF_C4steroid; 1.
DR PROSITE: PS00031; NUCLEAR RECEPTOR; UNKNOWN_1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 505 AA; 57624 MW; ACE6A50AE8AF694A CRC64;

Query Match 63.1%; Score 2212; DB 6; Length 505;
Best Local Similarity 86.7%; Pred. No. 6.8e-131;
Matches 436; Conservative 3; Mismatches 2; Indels 62; Gaps 1;
QY 1 MGETLGSPIDPESDSFTDLSANISQEMTWVDEMPFWPTNFGISSVDLSMEDHSHSF 60

```

|||||
Db 1 MGEETGDSPIDESDSFTTTLTANISOETIWDTEPWFPTNFGISSVDLSVMDHSHSF 60
Qy 61 DKKPPTVDFESSITPHYEDIPFTRTDPVADYKYDYLKOEYOSAIKVPASPPYSEKT 120
Db 61 DKKPPTVDFESSISAPHEDIPFTRTDPVADYKYDYLKOEYOSAIKVPASPPYSEKT 120
Qy 121 QLYN-----RNNCOYRFOKCLAVGSHNAI 124
Db 121 QLYNPHPEPNSLMAIECRVCGDKASGFHVGVAHEGCKGFRRRTIRLKLTYRCDLNC 180
Qy 125 -----RNNCOYRFOKCLAVGSHNAIRFGMRPOAEKKLLAEISSDIDQLNPSADLR 178
Db 181 RIRKRSRNNCOYRFOKCLAVGSHNAIRFGMRPOAEKKLLAEISSDIDQLNPSADLR 240
Qy 179 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPVIYDMNSLMGDEKIKFKHITPLOE 238
Db 241 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPVIYDMNSLMGDEKIKFKHITPLOE 300
Qy 239 OSKEVAIRIFOGCQFRSVEAVOETETAKSIRGFVNLDLNDQVTLTKYGVHEIITYMLAS 298
Db 301 OSKEVAIRIFOGCQFRSVEAVOETETAKSIRGFVNLDLNDQVTLTKYGVHEIITYMLAS 360
Qy 299 LMKKGVLSIEGQGFMTREFLSLKRPGDEMEPKFEPAVKNNALBLDSDLAIFIAVII 358
Db 361 LMKKGVLSIEGQGFMTREFLSLKRPGDEMEPKFEPAVKNNALBLDSDLAIFIAVII 420
Qy 359 LSGDRPGLLNVPRIEDIDONLQALELOLKLNPRESSOLFALLOKMTDLROIVTEHVOL 418
Db 421 LSGDRPGLLNVPRIEDIDONLQALELOLKLNPRESSOLFALLOKMTDLROIVTEHVOL 480
Qy 419 LOVIKKTETDMSLHPLLOEYKDL 443
Db 481 LOVIKKTETDMSLHPLLOEYKDL 505

RESULT 2
096J12 PRELIMINARY: PRT: 477 AA.
AC 096J12:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Unknown (protein for MGC:5041).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL; BC006811; AA06811.1; -.
DR InterPro: IPR000536; Hormone_rec.1lg.
DR InterPro: IPR001628; Znf_Casteroid.
DR Pfam; PF00104; hormone_rec.1.
DR Pfam; PF00105; zf-C4; 1.
DR ProDom; PD000035; Znf_Casteroid; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; UNKNOWN_1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 477 AA: 54681 MW: 1061C2074B739E0A CRC64:
Query Match 59.5%; Score 2089; DB 4; Length 477;
Best Local Similarity 87.0%; Pred.No.3.2e-12;
Matches 415; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

```

```

Qy 89 VVADYKYDYLKOEYOSAIKVPASPPYSEKTOLYN----- 124
Db 61 VVADYKYDYLKOEYOSAIKVPASPPYSEKTOLYNKHPHEPNSLMAIECRVCGDKASG 120
Qy 125 -----RNNCOYRFOKCLAVGSHNAI 146
Db 121 FHYGVACGCKGFRRRTIRLKLTYRCDLNCIRKRSRNNCOYRFOKCLAVGSHNAI 180
Qy 147 RGRMRPOAEKKLLAEISSDIDQLNPSADLRALKHLVDSYIKSFPLTKAKARAILTGK 206
Db 181 RGRMRPOAEKKLLAEISSDIDQLNPSADLRALKHLVDSYIKSFPLTKAKARAILTGK 240
Qy 207 TTDKSPVIYDMNSLMGDEKIKFKHITPLOSKEVAIRIFOGCQFRSVEAVOETETEA 266
Db 241 TTDKSPVIYDMNSLMGDEKIKFKHITPLOSKEVAIRIFOGCQFRSVEAVOETETEA 300
Qy 267 KSIPGFVNLDLNDQVTLTKYGVHEIITYMLASLMKKGVLISGQGFMTREFLSLRPF 326
Db 301 KSIPGFVNLDLNDQVTLTKYGVHEIITYMLASLMKKGVLISGQGFMTREFLSLRPF 360
Qy 327 GDFMEPKFEPAVKNNALBLDSDLAIFIAVIIISGDRPGLLNVPRIEDIDONLQALELO 386
Db 361 GDFMEPKFEPAVKNNALBLDSDLAIFIAVIIISGDRPGLLNVPRIEDIDONLQALELO 420
Qy 387 LKLNPESSOLFALLOKMTDLROIVTEHVOLQYIKKTETDMSLHPLLOEYKDL 443
Db 421 LKLNPESSOLFALLOKMTDLROIVTEHVOLQYIKKTETDMSLHPLLOEYKDL 477

RESULT 3
095J17 PRELIMINARY: PRT: 475 AA.
AC 095J17:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Peroxisome proliferator-activated receptor gamma 3 (peroxisome
DE proliferator-activated receptor gamma 1).
GN PPARGAMM3 OR PPARGAMM1.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhou J., Wilson K.M., Medh J.D.;
RT "Identification of novel peroxisome proliferator-activated receptor-
RT gamma splice variants and induction of PPAR-gamma expression by a
RT high-fat diet in monkey macrophages."
RL Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL; AY048696; AAL05263.1; -.
DR EMBL; AY048694; AAL05263.1; -.
DR InterPro: IPR000536; Hormone_rec.1lg.
DR InterPro: IPR001628; Znf_Casteroid.
DR Pfam; PF00104; hormone_rec.1.
DR Pfam; PF00105; zf-C4; 1.
DR ProDom; PD000035; Znf_Casteroid; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; UNKNOWN_1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 475 AA: 54470 MW: D16C3A931EB89181 CRC64:
Query Match 58.8%; Score 2064; DB 6; Length 475;
Best Local Similarity 86.1%; Pred.No.1.2e-12;
Matches 409; Conservative 2; Mismatches 2; Indels 62; Gaps 1;

```

```

Qy 29 MNVDTMPFWPTNFGISSVDLSVMDHSHSFDIKPPTVDFESSISAPHEDIPFTRTDPV 88
Db 1 MNVDTMPFWPTNFGISSVDLSVMDHSHSFDIKPPTVDFESSISAPHEDIPFTRTDPV 60

```

```

Qy 31 MVDTEMPFWPTNFGISSVDLSVMDHSHSFDIKPPTVDFESSISAPHEDIPFTRTDPV 90
Db 1 MVDTEMPFWPTNFGISSVDLSVMDHSHSFDIKPPTVDFESSISAPHEDIPFTRTDPV 60

```



```

QY 91 ADYKDKLQEOYSAIKVEPASPPIYSEKTOLYN-----124
D6 61 ADYKDKLQEOYSAIKVEPASPPIYSEKTOLYNKPHEPSPNSLMAIBCRVCGDKASGFH 120
QY 125 -----RNNCOYCRFOKCLAVGMSHNAIRF 148
D6 121 YGVHACEGCKGFRRRTIRLKLHYDRCDLNCRIHKSSRNKCOYCRFOKCLAVGMSHNAIRF 180
QY 149 GMPQAEKELIAETSSIDOLNPESADLRALAKHLVSYIKSPFLTKAKARAILTGKTT 208
D6 161 GMPQAEKELIAETSSIDOLNPESADLRALAKHLVSYIKSPFLTKAKARAILTGKTT 240
QY 209 DKSPFVIYDMNSLMNGEDKIKFKHTTLPLOESKEVAIRIFOGCQFRSVEAVOEITEYAKS 268
D6 241 DKSPFVIYDMNSLMNGEDKIKFKHTTLPLOESKEVAIRIFOGCQFRSVEAVOEITEYAKS 300
QY 269 IPGFVNLNDQVTLTKYGVHEIITYTMLASLMNKDGVLSISGQGFMTREFLSLRKPPGD 328
D6 301 IPGFVNLNDQVTLTKYGVHEIITYTMLASLMNKDGVLSISGQGFMTREFLSLRKPPGD 360
QY 329 FMEPKFEFAVKFNALIEDSDLAIFAVIILSGDRPGLNKPREDIODNLQALELQLK 388
D6 361 FMEPKFEFAVKFNALIEDSDLAIFAVIILSGDRPGLNKPREDIODNLQALELQLK 420
QY 389 LNHPESSOLFVKLQKMTDLRQIYTEHVQOLLQVIRKTTDMSLHPLOEIKKDL 443
D6 421 LNHPESSOLFVKLQKMTDLRQIYTEHVQOLLQVIRKTTDMSLHPLOEIKKDL 475

```

RESULT 4

099PC7 PRELIMINARY; PRT: 475 AA.

```

ID 099PC7 PRELIMINARY; PRT: 475 AA.
AC 099PC7:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Peroxisome proliferator activated receptor-gamma (Fragment).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Hystriognathii; Cavidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADIPOSE TISSUE;
RA Khoo B.Y., Najimudin N., Tengku Muhammad T.S.:
RT "Molecular Cloning and Sequencing of Peroxisome Proliferator Activated
RT Receptor-gamma (PPARGamma) from Guinea Pig.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL: AF317514; AAG60685.2; -
DR HSSP: P37231; 1A99.
DR InterPro: IPR000536; Hormone_rec_1ig.
DR InterPro: IPR001723; Stdhnm_receptor.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF00104; hormone_rec_1.
DR PRINTS: PR00398; STPDHORMER.
DR PRINTS: PR00047; STROIDFINGER.
DR PRODom: PD000035; Znf_C4steroid.1.
DR SMART: SM00430; HOL1; 1.
DR SMART: SM00399; Znf_C4; 1.
DR PROSITE: PS00031; NUCLEAR RECEPTOR; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 475 AA; 54425 MW; DE45CF49C896F401 CRC64;

```

Query Match 58.3%; Score 2045; DB 11; Length 475;
 Best Local Similarity 85.1%; Pred. No. 1.9e-120;
 Matches 404; Conservative 5; Mismatches 4; Indels 62; Gaps 1;

```

QY 31 MVDTEMFPWPNFGISSVDLSVMEDSHSHFDIKPTTVDFSSISTPHYEDIPTRTDPV 90
D6 1 MVDTEMFPWPNFGISSVDLSVMEDSHSHFDIKPTTVDFSSISAPRHEDIPTRADPMV 60
QY 91 ADYKDKLQEOYSAIKVEPASPPIYSEKTOLYN-----124
D6 61 ADYKDKLQEOYSAIKVEPASPPIYSEKTOLYNKPHEPSPNSLMAIBCRVCGDKASGFH 120
QY 125 -----RNNCOYCRFOKCLAVGMSHNAIRF 148
D6 121 YGVHACEGCKGFRRRTIRLKLHYDRCDLNCRIHKSSRNKCOYCRFOKCLAVGMSHNAIRF 180
QY 149 GMPQAEKELIAETSSIDOLNPESADLRALAKHLVSYIKSPFLTKAKARAILTGKTT 208
D6 161 GMPQAEKELIAETSSIDOLNPESADLRALAKHLVSYIKSPFLTKAKARAILTGKTT 240
QY 209 DKSPFVIYDMNSLMNGEDKIKFKHTTLPLOESKEVAIRIFOGCQFRSVEAVOEITEYAKS 268
D6 241 DKSPFVIYDMNSLMNGEDKIKFKHTTLPLOESKEVAIRIFOGCQFRSVEAVOEITEYAKS 300
QY 269 IPGFVNLNDQVTLTKYGVHEIITYTMLASLMNKDGVLSISGQGFMTREFLSLRKPPGD 328
D6 301 IPGFVNLNDQVTLTKYGVHEIITYTMLASLMNKDGVLSISGQGFMTREFLSLRKPPGD 360
QY 329 FMEPKFEFAVKFNALIEDSDLAIFAVIILSGDRPGLNKPREDIODNLQALELQLK 388
D6 361 FMEPKFEFAVKFNALIEDSDLAIFAVIILSGDRPGLNKPREDIODNLQALELQLK 420
QY 389 LNHPESSOLFVKLQKMTDLRQIYTEHVQOLLQVIRKTTDMSLHPLOEIKKDL 443
D6 421 LNHPESSOLFVKLQKMTDLRQIYTEHVQOLLQVIRKTTDMSLHPLOEIKKDL 475

```

RESULT 5

091878 PRELIMINARY; PRT: 475 AA.

```

ID 091878 PRELIMINARY; PRT: 475 AA.
AC 091878:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Peroxisome proliferator-activated receptor (Peroxisome
DE proliferator-activated receptor gamma).
OS Gallus gallus (Chicken).
GN PPAR GAMMA.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato K., Akiba Y.;
RT "Characterization of a cDNA sequence encoding the peroxisome
RT proliferator activated receptor gamma in broiler chicken.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC [2]
CC SEQUENCE FROM N.A.
RX MEDLINE=20267232; PubMed=10809235;
RA Takada I., Yu R.T., Xu H.E., Lambert M.H., Montana V.G., Klierer S.A.,
RA Evans R.M., Umesono K.;
RT "Alteration of a single amino acid in peroxisome proliferator-
RT activated receptor-alpha (PPAR alpha) in peroxisome proliferator-
RT phenotype.";
RL Mol. Endocrinol. 14:733-740(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RA Meng H., Li J., Zhao J., Wang Q., Li H.;
RT "The SNPs detection on exon 2 of PPAR gamma gene in chicken.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL: AB045597; BAA98100.1; -
DR EMBL: AF163811; AAF80170.1; -

```

DR EMBL: AF470456; AAL85323.1; -.
 DR HSSP: P37231; 1FM9.
 DR InterPro: IPR000536; Hormone_rec_11g.
 DR InterPro: IPR001723; Stdhmn_receptor.
 DR InterPro: IPR001628; ZnF_Csteroi.
 DR Pfam: PF00104; hormone_rec.1.
 DR PRINTS: PR00105; zf-C4; 1.
 DR PRINTS: PR00398; STRDHOMNER.
 DR PRINTS: PR00047; STROIDFINGER.
 DR ProDom: PD000035; ZnF_Csteroi.1.
 DR SMART: SM00430; HOLI; 1.
 DR SMART: SM00399; ZnF_C4; 1.
 DR PROSITE: PS00031; NOCLEAR_RECEPTOR; 1.
 DR DNA-binding: Nuclear protein; Receptor; Transcription regulation;
 zinc-finger.
 KW ZINC_FINGER.
 SQ SEQUENCE 475 AA; 54445 MW; 8CB1C23898E2C5D5 CRC64;

Query Match 54.7%; Score 1920; DB 13; Length 475;
 Best Local Similarity 80.0%; Pred. No. 1.3e-112;
 Matches 380; Conservative 14; Mismatches 19; Indels 62; Gaps 1;

QY 31 MVDTEMPMPNTEGSSVDSLVSMEHSHSFDIKPTVDSSISPHYEDIPFTTDPVY 90
 DB 1 MVDTEMPMPNTEGSSVDSLVSMEHSHSFDIKPTVDSSISPHYEDIPFTTDPVY 90
 QY 91 ADVKDKLQEOYSALKEYPASPPYSEKTOLYN----- 124
 DB 61 IDYKDKLQDQSAIKMEPPSPYFSEKVOYLYNKPHESSNSLMAIECRVCGDKASGFH 120
 QY 125 -----RNKCYCRQKCLAVGSHNAIRF 148
 DB 121 YGVHACEGCKGFRRRTIRKLTYDRCDLNCRIHKKS RNKCYCRQKCLAVGSHNAIRF 180
 QY 149 GMPQAEKELIAETSSDIDLPESADLRALAKHLYSYIKSPFLTKAKARAILTGKTT 208
 DB 181 GMPQAEKELIAETSSDIDLPESADLRALAKHLYSYIKSPFLTKAKARAILTGKTT 240
 QY 209 DKSPFVIYDMNSLMNGEDKIKFKHTPLQEOSEKVAIRIFQCGCFRSYEAVOETIEYAKS 268
 DB 241 DKSPFVIYDMNSLMNGEDKIKFKHTPLQEOSEKVAIRIFQCGCFRSYEAVOETIEYAKS 300
 QY 269 IPGFVNLDLNDQVTLTKYGVHEIITYTMLASLMNKDGVLSISGCGFMTREPLKSLRKPFGD 328
 DB 301 IPGFVNLDLNDQVTLTKYGVHEIITYTMLASLMNKDGVLSISGCGFMTREPLKSLRKPFGD 360
 QY 329 FMEPKFEFAVKFNALELDSDLAIFIAVILISGDRPGLLNKVPRIEDIDNLLQALELQK 388
 DB 361 FMEPKFEFAVKFNALELDSDLAIFIAVILISGDRPGLLNKVPRIEDIDNLLQALELQK 420
 QY 389 LNHPESSQLFAKLQKMTDLRQIVTEHVQLQVIRKTTEDMSLHPLQEIYK 440
 DB 421 LNHPESSQLFAKLQKMTDLRQIVTEHVQLQVIRKTTEDMSLHPLQEIYK 475

RESULT 6

Q80GCO PRELIMINARY; PRT; 472 AA.
 AC 080GCO:
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Peroxisome proliferator-activated receptor gamma (Fragment).
 GN PPARGAMA.
 OS Anser anser (domestic goose).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Anser.
 OX NCBI_taxid=8843;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE-FAT;
 RA Meng H., Li H.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF481798; AAL87700.1; -.

KW Receptor.
 FT NON_TER 472 472
 SQ SEQUENCE 472 AA; 54095 MW; CD17ED2FF8FAFD25 CRC64;

Query Match 54.6%; Score 1916; DB 13; Length 472;
 Best Local Similarity 80.1%; Pred. No. 2.3e-112;
 Matches 378; Conservative 16; Mismatches 16; Indels 62; Gaps 1;

QY 31 MVDTEMPMPNTEGSSVDSLVSMEHSHSFDIKPTVDSSISPHYEDIPFTTDPVY 90
 DB 1 MVDTEMPMPNTEGSSVDSLVSMEHSHSFDIKPTVDSSISPHYEDIPFTTDPVY 90
 QY 91 ADVKDKLQEOYSALKEYPASPPYSEKTOLYN----- 124
 DB 61 IDYKDKLQDQSAIKMEPPSPYFSEKVOYLYNKPHESSNSLMAIECRVCGDKASGFH 120
 QY 125 -----RNKCYCRQKCLAVGSHNAIRF 148
 DB 121 YGVHACEGCKGFRRRTIRKLTYDRCDLNCRIHKKS RNKCYCRQKCLAVGSHNAIRF 180
 QY 149 GMPQAEKELIAETSSDIDLPESADLRALAKHLYSYIKSPFLTKAKARAILTGKTT 208
 DB 181 GMPQAEKELIAETSSDIDLPESADLRALAKHLYSYIKSPFLTKAKARAILTGKTT 240
 QY 209 DKSPFVIYDMNSLMNGEDKIKFKHTPLQEOSEKVAIRIFQCGCFRSYEAVOETIEYAKS 268
 DB 241 DKSPFVIYDMNSLMNGEDKIKFKHTPLQEOSEKVAIRIFQCGCFRSYEAVOETIEYAKS 300
 QY 269 IPGFVNLDLNDQVTLTKYGVHEIITYTMLASLMNKDGVLSISGCGFMTREPLKSLRKPFGD 328
 DB 301 IPGFVNLDLNDQVTLTKYGVHEIITYTMLASLMNKDGVLSISGCGFMTREPLKSLRKPFGD 360
 QY 329 FMEPKFEFAVKFNALELDSDLAIFIAVILISGDRPGLLNKVPRIEDIDNLLQALELQK 388
 DB 361 FMEPKFEFAVKFNALELDSDLAIFIAVILISGDRPGLLNKVPRIEDIDNLLQALELQK 420
 QY 389 LNHPESSQLFAKLQKMTDLRQIVTEHVQLQVIRKTTEDMSLHPLQEIYK 440
 DB 421 LNHPESSQLFAKLQKMTDLRQIVTEHVQLQVIRKTTEDMSLHPLQEIYK 472

RESULT 7

Q98SF8 PRELIMINARY; PRT; 477 AA.
 AC 098SF8:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Peroxisome proliferator-activated receptor gamma.
 GN NR1C3.
 OS Xenopus laevis (African clawed frog).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_taxid=8355;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE-FAT BODY;
 RA Escher P.;
 RT "Structural and functional analysis of PPARs."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
 DR EMBL: AJ310087; CAC34402.1; -.
 DR HSSP: P37231; 1FM9.
 DR InterPro: IPR000536; Hormone_rec_11g.
 DR InterPro: IPR001723; Stdhmn_receptor.
 DR InterPro: IPR001628; ZnF_Csteroi.
 DR Pfam: PF00104; hormone_rec.1.
 DR Pfam: PF00105; zf-C4; 1.
 DR PRINTS: PR00398; STRDHOMNER.
 DR PRINTS: PR00047; STROIDFINGER.
 DR ProDom: PD000035; ZnF_Csteroi.1.

DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR000398; STRDHOMONER.
DR PRINTS; PR000447; STROIDFINGER.
DR Prodom; PD000035; znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
DR DNA-binding; Nuclear protein; Receptor; Transcription regulation;
zinc-finger.
KW NON_TER 1 1
FT NON_TER 393 393
SQ SEQUENCE 393 AA; 44812 MW; 115E7D61AAB19E6F CRC64;

Query Match 44.2%; Score 1549; DB 13; Length 393;
Best Local Similarity 79.6%; Pred. No. 1.8e-89;
Matches 313; Conservative 9; Mismatches 9; Indels 62; Gaps 1;

OY 105 AIKVPASPPYSEKTOLYN-----RNKCYCRFOKCLAVGSHNAIRFGMRPOAEKELAE 124
DB 1 AIKMEPPSPYFSEKVOLYNKPPEETNSLMAIEGRVCGDKASGFHYGVHACEGCKGEFR 60
OY 125 -----RNKCYCRFOKCLAVGSHNAIRFGMRPOAEKELAE 162
DB 61 RTIRKLTYRCDLNCRIHKSRNKCQYCRFOKCLAVGSHNAIRFGMRPOAEKELAE 120
OY 163 ISSDIDOLNPESADLRALAKHLYDSYIKSFPLTKAKARAILGKTTDRSPFVIYDMSLM 222
DB 121 ISSDIDOLNPESADLRALAKHLYDSYIKSFPLTKAKARAILGKTTDRSPFVIYDMSLM 180
OY 223 MGEKIKRKHITPLOBOSKVAIRIFGCGQFRSVAVOETITAYASIGFVNLNDQYT 282
DB 223 MGEKIKRKHITPLOBOSKVAIRIFGCGQFRSVAVOETITAYASIGFVNLNDQYT 240
OY 181 MEEDIDINKHVTPLDOKNEVAIRIFORCFRSVAVOETITAFASIPFVNPDLNDQYT 240
DB 283 LKTYVHEITITMLASLNKMDGVLSSEGQFMTREFLSLRKPFDFMPKREFPAVKENA 342
OY 241 LKTYVHEITITMLASLNKMDGVLSSEGQFMTREFLSLRKPFDFMPKREFPAVKENA 300
DB 241 LKTYVHEITITMLASLNKMDGVLSSEGQFMTREFLSLRKPFDFMPKREFPAVKENA 300
OY 343 LELDDSDLAIFLAVIILSGDRPGLLNKPIEDIDQNLQALELQKLNHPSSQLFAKLL 402
DB 301 LELDDSDLAIFLAVIILSGDRPGLLNKPIEDIDQNLQALELQKLNHPSSQLFAKLL 360
OY 403 QKMTDLROIVTEHVOLLQYIKKTEIDMSLHPLL 435
DB 361 QKMTDLROIVTEHVOLLQYIKKTEIDMSLHPLL 393

RESULT 10
O9PM01 PRELIMINARY: PRT; 533 AA.
AC O9PM01; 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Peroxisome proliferator-activated receptor gamma.
GN PPAR.
OS Pleuronectes platea (Plaice).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Pleuronectidae; Pleuronectes.
OX NCBI_TaxId=8262;
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE=Gill;
RA Loefer M.J.;
RT "A peroxisomal proliferator-activated receptor gamma gene from a
marine fish."
RL Submitted (AUG-1956) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL; AJ243956; CAB51618.1; -.
DR HSRP; P37231; 1FW9.
DR InterPro; IPR000536; Hormone_rec_1lg.

DR InterPro; IPR001723; Stdhmn_receptor.
DR InterPro; IPR001628; znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR000398; STRDHOMONER.
DR Prodom; PD000035; znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
DR DNA-binding; Nuclear protein; Receptor; Transcription regulation;
zinc-finger.
KW NON_TER 1 1
FT NON_TER 393 393
SQ SEQUENCE 393 AA; 60513 MW; DD395E777961300 CRC64;

Query Match 37.8%; Score 1325; DB 13; Length 533;
Best Local Similarity 50.6%; Pred. No. 3e-75;
Matches 276; Conservative 59; Mismatches 66; Indels 144; Gaps 6;

OY 31 NYDT-EMPPWPTNFGISSVDLSVMEDSHSPDIKFTTVDESSIST----- 75
DB 1 NYDTQOLLFWPVGFSLASVDLSLDDSSHSLDKHILATIDYTSISASVPSLSFQLMSS 60
OY 76 -----PHYEDIPFT-----RTDPVYADYKTKDQEQSAIKVPASPP 114
DB 61 ISSVGMAVDPSPQSEEHLLNMDYTNMHSYRTEPV-----HNSIKMEPESP 108
OY 115 YNSE----- 118
DB 109 QYSDPVSFSKLDDDTTASLNIECRVCGDKASGFHYGVHACEGCKGFPRTIRKLVDH 168
OY 119 -----KTOLYNBKCQYCRFOKCLAVGSHNAIRFGMRPOAEKELAEISSDIDOLNP 173
DB 169 CDLHCRIRHKSRNKCQYCRFOKCLAVGSHNAIRFGMRPOAEKELAEISSDIDOLNP 228
OY 174 SADLRALAKHLYDSYIKSFPLTKAKARAILGKTTDRSPFVIYDMSLMGDKIFKHI 233
DB 228 SADLRALAKHLYDSYIKSFPLTKAKARAILGKTTDRSPFVIYDMSLMGDKIFKHI 288
OY 234 TPLDQSK-----EVAIRIFGCGQFRSV 258
DB 269 MPEQDQASVLTATNRGLTEHHMGSDYGVWGTSTISGQEPONALELRFPOCQSRSA 348
OY 318 VOETIEYAKSITGFVNLNDQYTLLKTYVHEITITMLASLNKMDGVLSSEGQFMTREF 318
DB 349 VREVTETAFKSIPEFTDLDNDQYTLKTYVHEITITMLASLNKMDGVLSSEGQFMTREF 408
OY 319 LKSLRPGDFMEKREFPAVKENALELDDSDLAIFLAVIILSGDRPGLLNKPIEDIDON 378
DB 409 LKSLRPGDFMEKREFPAVKENALELDDSDLAIFLAVIILSGDRPGLLNKPIEDIDON 468
OY 438 LLOALELQKLNHPSSQLFAKLLQKMTDLROIVTEHVOLLQYIKKTEIDMSLHPLLQEI 438
DB 469 VLHSELQKLNHPSSQLFAKLLQKMTDLROIVTEHVOLLQYIKKTEIDMSLHPLLQEI 528
OY 439 YKDL 443
DB 529 MKDLY 533

RESULT 11
O9W712 PRELIMINARY: PRT; 532 AA.
AC O9W712; 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Peroxisome proliferator-activated receptor gamma.
GN PPAR.
OS Pleuronectes platea (Plaice).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;
OC Pleuronectoidae; Pleuronectidae; Pleuronectes.

OY 439 YKDL 443
DB 539 MRDL 543

RESULT 13

OY18W3 PRELIMINARY; PRT; 443 AA.
ID OY18W3
AC OY18W3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Peroxisome proliferator-activated receptor beta.
GN PPARBETA.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauiria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=20267232; PubMed=10809235;
RA Takada I., Yu R.T., Xu H.E., Lambert M.H., Montana V.G., Kilewer S.A.,
RA Evans R.M., Umeson K.;
RT "Alteration of a single amino acid in peroxisome proliferator-
RT activated receptor-alpha (PPAR alpha) generates a PPAR delta
RT phenotype.";
RT Mol. Endocrinol. 14:733-740(2000).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL; AF163810; AAF80480.1; -.
DR HSSP; P03372; IHCQ.
DR InterPro; IPR000536; Hormone_rec.1lg.
DR InterPro; IPR001723; Sterm_receptor.
DR InterPro; IPR001628; znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00398; STRDHOMER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRODOM; PD000035; znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 443 AA; 50108 MW; EB6E0C39554C76CD CQC64;

Query Match 31.9%; Score 1117.5; DB 13; Length 443;
Best Local Similarity 53.2%; Pred. No. 2.4e-62;
Matches 235; Conservative 68; Mismatches 88; Indels 51; Gaps 10;

OY 14 SDSFDD--TISANISOEMTVDTMPFMPNFGISSVDLSVMEDSHSFDTKPTTYDF 70
DB 41 SSSYTDLSOSSSPSISDLOMGCE-----TASGALNVECRVCGKASGF----- 85
OY 71 SSISSPHY-----EDIP--FTRTDPVADYKYLKLOEYQSAIKVEPASPYYSEKTOLY 123
DB 86 -----HYGVHACGECCKGFERR-----IRMKL-EYEKER-----SCKIKQK 121
OY 124 NNNKOYGRFOGCLAVGSHNAIRGRMPQAEKELIAET-SSDIDQLNPESADIRALAK 182
DB 122 NNNKOYGRFOGCLGSHNAIRGRMPEAKERKLVAGLTASETISCONPOVADVAKESK 181
OY 183 HLYDSYKSPFLTKAKARAILTGK--TTDKSPVIYDMNSLMNGEDKIKFKHTTPILOESK 241
DB 182 HIYNAIYNFNMTTKKAKGILTKGASSTPQPFVHDMDTLMOAEGVLWKLVNCTIPPYK 241
OY 242 EVAIRIFOGCFRSEVAEOETEVAKSIPGFVNLNDQVTLTKYGVHEIITMLASLNN 301
DB 242 EIGVHVFRCCQTVEYVRELTEFAKSIPIFGILNDQVTLTKYGVHEIITMLASLNN 301
OY 302 KDGVLISGCGMTTEFLKSLKPPGDMPEKFEFAVKFNALELDDSLAIFAVIIISG 361
DB 302 KDGVLISGCGMTTEFLKSLKPPGDMPEKFEFAVKFNALELDDSLAIFAVIIISG 361

DB 302 KDGVLVANGNGVFREFLTLLKPNEMEPEKFEFAVKFNALELDDSLFVAAILCG 361
OY 362 DRPGLLVNVPIDIDONLQALELOLKLNHPESSOLFALOKMTDLROIVTEHVOLOV 421
DB 362 DRPGLLVNVPIDIDONLQALELFEHLQSNHPDQYLFKLLQKMDRLQVTEHQAOLVK 421
OY 422 IKKTETMSLHPLQEIYKDL 443
DB 422 IKKTETETSLHPLQEIYKDMY 443

RESULT 14

OY9SK24 PRELIMINARY; PRT; 219 AA.
ID OY9SK24
AC OY9SK24;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Peroxisome proliferator-activated receptor gamma (Fragment).
OS Mustela vison (American mink).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Mustelidae; Mustelinae;
OC Mustela.
OX NCBI_TaxID=9667;
RN (1)
RP SEQUENCE FROM N.A.
RX Desmarais J.A., Bennett R.D., Murphy B.D.;
RA "Mustela vison Peroxisome Proliferator-Activated Receptor Gamma
RT MRNA.";
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AY050524; AAL14452.1; -.
DR InterPro; IPR000536; Hormone_rec.1lg.
DR InterPro; IPR001628; znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; zf-C4; 1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
FT NON_TER 1 1
FT NON_TER 219 219
SQ SEQUENCE 219 AA; 24828 MW; 8F1D1E4F90EA6320 CQC64;

Query Match 31.5%; Score 1105; DB 6; Length 219;
Best Local Similarity 99.5%; Pred. No. 5.7e-62;
Matches 218; Conservative 99.5%; 1; Mismatches 0; Indels 0; Gaps 0;

OY 136 CLAVGMSHNAIRFGMPQAEKELIAETSSDIDQLNPESADIRALAKHLVDSYKSPFLT 195
DB 1 CLAVGMSHNAIRFGMPQAEKELIAETSSDIDQLNPESADIRALAKHLVDSYKSPFLT 60
OY 196 KAKARAILTGKTTDKSPVIYDMNSLMNGEDKIKFKHTTPILOESKEVAIRIFOGCGRS 255
DB 61 KAKARAILTGKTTDKSPVIYDMNSLMNGEDKIKFKHTTPILOESKEVAIRIFOGCGRS 120
OY 256 VEAVOETEVAKSIPGFVNLNDQVTLTKYGVHEIITMLASLNNKGVLSISGCGMT 315
DB 121 VEAVOETEVAKNIPGFVNLNDQVTLTKYGVHEIITMLASLNNKGVLSISGCGMT 180
OY 316 REFLKSLKPPGDMPEKFEFAVKFNALELDDSLAIFI 354
DB 181 REFLKSLKPPGDMPEKFEFAVKFNALELDDSLAIFI 219

RESULT 15

OY9SN78 PRELIMINARY; PRT; 468 AA.
ID OY9SN78
AC OY9SN78;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Peroxisome proliferator activated receptor alpha.
OS Canis familiaris (Dog).

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2003, 03:23:21 ; Search time 20 seconds
(without alignments)
1426.786 Million cell updates/sec

Title: US-09-931-007A-1
Perfect score: 3508
Sequence: 1 MGELTGDSPIDPESDSFTDT.....KTEIDMSLPLQLQEIYKDLV 688

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2231	63.6	505	1 PPAT_HUMAN	P37231 homo sapien
2	2216	63.2	505	1 PPAT_MACMU	O18824 macaca mula
3	2171.5	61.9	504	1 PPAT_PIG	O62807 sus scrofa
4	2153	61.4	505	1 PPAT_MOUSE	P37238 mus musculu
5	2148	61.2	505	1 PPAT_RAT	O88275 rattus norv
6	2132	60.8	505	1 PPAT_BOVIN	O18971 bos taurus
7	2043	58.2	475	1 PPAT_CRICR	P57297 cricetus
8	2041	58.2	475	1 PPAT_XENLA	O19052 oryctolagus
9	1471	41.9	477	1 PPAT_XENLA	P37234 xenopus lae
10	1101.5	31.4	441	1 PPAS_HUMAN	Q03181 homo sapien
11	1091.5	31.1	468	1 PPAS_HUMAN	P37230 rattus norv
12	1083.5	30.9	468	1 PPAS_MOUSE	P23204 mus musculu
13	1082.5	30.9	440	1 PPAS_MOUSE	P33596 mus musculu
14	1076.5	30.7	468	1 PPAS_HUMAN	Q07869 homo sapien
15	1060.5	30.2	474	1 PPAS_XENLA	P37232 xenopus lae
16	1042	29.7	467	1 PPAS_CAVPO	O35507 cavia porce
17	946.5	27.0	396	1 PPAS_XENLA	P37233 xenopus lae
18	431	12.3	606	1 E75_METEN	O60674 mus musculu
19	408	11.6	576	1 NR02_MOUSE	O63504 rattus norv
20	404.5	11.5	579	1 NR02_HUMAN	O14995 homo sapien
21	397	11.3	578	1 NR02_RAT	O63504 rattus norv
22	392	11.2	1237	1 E75A_DROME	P17672 drosophila
23	385.5	11.0	1237	1 E75A_DROME	P17672 drosophila
24	379	10.8	614	1 NR01_HUMAN	P20393 homo sapien
25	375.5	10.7	459	1 R0RB_RAT	P45446 rattus norv
26	371.5	10.6	508	1 NR01_RAT	O63503 rattus norv
27	371.5	10.6	1443	1 E75C_DROME	P13055 drosophila
28	365.5	10.4	459	1 R0RB_HUMAN	O92753 galliera me
29	354.5	10.1	711	1 E75_GALME	O08893 manduca sex
30	351	10.0	699	1 E75_MANSE	O01339 choriscoueu
31	344	9.8	690	1 E75_CHOFO	P33598 homo sapien
32	341.5	9.7	556	1 R0RA_MOUSE	P51448 mus musculu
33	338.5	9.6	523	1 R0RA_MOUSE	P51448 mus musculu

34	327.5	9.3	458	1 RRA_NOTVI	P18514 notophthalm
35	326.5	9.3	455	1 RRB_CHICK	P22448 gallus gall
36	326.5	9.3	455	1 RRB_CONJA	O9w6b3 coturnix co
37	323.6	9.3	447	1 RRA_PIGRU	O9w523 fugu rubrip
38	323.5	9.3	560	1 R0RG_HUMAN	P51449 homo sapien
39	324.5	9.3	455	1 RRB_HUMAN	P10826 homo sapien
40	323	9.2	865	1 E78A_DROME	P45447 drosophila
41	322	9.2	445	1 NR03_MOUSE	O920Y9 mus musculu
42	322	9.2	445	1 NR03_RAT	O62685 rattus norv
43	319.5	9.1	482	1 RRB_MOUSE	P22605 mus musculu
44	318	9.1	505	1 RRB_NOTVI	P18516 notophthalm
45	316	9.0	476	1 RRG1_XENLA	P51127 xenopus lae

ALIGNMENTS

RESULT 1
ID PPAT_HUMAN STANDARD; PRT; 505 AA.
AC P37231: Q14515; Q15178; Q15832; Q00684; Q15180; Q00710;
DT 01-OCT-1994 (Rel. 30, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).
GN PPARG OR NR1C3.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND CHARACTERIZATION.
RC TISSUE=Heart;
RX MEDLINE=97218249; PubMed=9065481;
RA Mukherjee R., Jow L., Croston G.E., Paterniti J.R. Jr.;
RT "Identification, characterization, and tissue distribution of human
RT peroxisome proliferator-activated receptor (PPAR) isoforms PPARgamma2
RT versus PPARgamma1 and activation with retinoid X receptor agonists and
RT antagonists";
RL J. Biol. Chem. 272:8071-8076(1997).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Fat body;
RA Elbrecht A., Chen Y., Cullinan C.A., Hayes N., Leibowitz M.D.,
RT Moller D.E., Berger J.;
RT "Molecular cloning, expression and characterization of human
RT peroxisome proliferator activated receptor gamma 1 and gamma 2.";
RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Adipose tissue;
RA Kato S.;
RT Submitted (FEB-1996) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Bone marrow;
RX MEDLINE=95307078; PubMed=7787419;
RA Greene M.E., Blumberg B., McBride O.W., Yi H.F., Kronquist K.,
RT Kwan K., Hsieh L., Greene G., Nimer S.D.;
RT "Isolation of the human peroxisome proliferator activated receptor
RT gamma cDNA: expression in hematopoietic cells and chromosomal
RT mapping.";
RL Gene Expr. 4:281-299(1995).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RX MEDLINE=98016122; PubMed=9356045;
RA Okazawa H., Mori H., Tamori Y., Araki S., Niki T., Masugi J.,
RT Kawachi M., Kubota T., Shinoda H., Kasuga M.;
RT "No coding mutations are detected in the peroxisome proliferator-
RT activated receptor-gamma gene in Japanese patients with lipotrophic
RT diabetes.";
RL Diabetes 46:1904-1906(1997).

RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Placenta.
 RX MEDLINE=96305359; PubMed=8706692;
 RA Lambe K.G., Tugwood J.D.;
 RT "A human peroxisome-proliferator-activated receptor-gamma is
 RT activated by inducers of adipogenesis, including thiazolidinedione
 RT drugs.";
 RL Eur. J. Biochem. 239:1-7(1996).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
 RX MEDLINE=20337987; PubMed=10882139;
 RA Gampe R.T. Jr., Montana V.G., Lambert M.H., Miller A.B., Bledsoe R.K.,
 RA Milburn M.V., Klierer S.A., Willson T.M., Xu H.E.;
 RT "Asymmetry in the PARGamma/RXRalpha crystal structure reveals the
 RT molecular basis of heterodimerization among nuclear receptors.";
 RL Mol. Cell 5:545-555(2000).
 RN [8]
 RP VARIANT ALA-12.
 RX MEDLINE=98086341; PubMed=9425261;
 RA Yen C.-J., Beamer B.A., Negri C., Silver K., Brown K.A., Varnall D.P.,
 RA Burns D.K., Roth J., Shuldiner A.R.;
 RT "Molecular scanning of the human peroxisome proliferator activated
 RT receptor gamma (hPPAR-gamma) gene in diabetic Caucasians:
 RT identification of a pro22ala PPAR-gamma-2 missense mutation.";
 RL Biochem. Biophys. Res. Commun. 241:270-274(1997).
 RN [9]
 RP VARIANT OBESITY GLN-113.
 RX MEDLINE=98418646; PubMed=9753710;
 RA Ristow M., Muller-Wieland D., Pfeiffer A., Krone W., Kahn C.R.;
 RT "Obesity associated with a mutation in a genetic regulator of
 RT adipocyte differentiation.";
 RL New Engl. J. Med. 339:953-959(1998).
 RN [10]
 RP VARIANT ALA-12.
 RX MEDLINE=99337654; PubMed=10407229;
 RA Hamann A., Munzberg H., Buttron P., Busing B., Hinney A., Mayer H.,
 RA Siegfried W., Hebebrand J., Grotten H.;
 RT "Missense variants in the human peroxisome proliferator-activated
 RT receptor gamma2 gene in lean and obese subjects.";
 RL Eur. J. Endocrinol. 141:90-92(1999).
 RN [11]
 RP VARIANTS COLON CANCER PRO-314 AND HIS-316, AND VARIANT ALA-12.
 RX MEDLINE=99322672; PubMed=10394368;
 RA Sarraf P., Mueller E., Smith W.M., Wright H.M., Kum J.B.,
 RA Aaltonen L.A., de la Chapelle A., Spiegelman B.M., Eng C.;
 RT "Loss-of-function mutations in PPAR-gamma associated with human colon
 RT cancer.";
 RL Mol. Biol. Cell 3:799-804(1999).
 RN [12]
 RP VARIANTS DIABETES MET-318 AND LEU-495.
 RX MEDLINE=20085964; PubMed=10622252;
 RA Barroso I., Gurnell M., Crowley V.E.F., Agostini M., Schwabel J.W.,
 RA Soos M.A., Masien G.L., Williams T.D.M., Lewis H., Schaefer A.J.,
 RA Chatterjee V.K.K., O'Rahilly S.;
 RT "Dominant negative mutations in human PPAR-gamma associated with
 RT severe insulin resistance, diabetes mellitus and hypertension.";
 RL Nature 402:880-883(1999).
 CC -1- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
 CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
 CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
 CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
 CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
 CC OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.
 CC -1- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Nucleus.
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 AND 2 (SHOWN HERE);
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADIPOSE TISSUE. LOWER IN
 CC SKELETAL MUSCLE, SPLEEN, HEART AND LIVER. ALSO ARE DETECTABLE IN
 CC PLACENTA, LUNG AND OVARY.
 CC -1- DISEASE: DEFECTS IN PARG CAN LEAD TO TYPE 2 INSULIN-RESISTANT
 CC DIABETES AND HYPERTENSION.

CC -1- DISEASE: DEFECTS IN PARG COULD PLAY A ROLE IN THE GENETIC
 CC PREDISPOSITION TO OBESITY.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NRI SUBFAMILY.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U79012; AAC51248.1; -;
 DR EMBL: U63415; AAB04028.1; -;
 DR EMBL: D83233; BAA18949.1; -;
 DR EMBL: LA0904; AAA80314.1; -;
 DR EMBL: AB005526; BAA23354.1; ALT_INIT.
 DR EMBL: AB005521; BAA23354.1; JOINED.
 DR EMBL: AB005522; BAA23354.1; JOINED.
 DR EMBL: AB005523; BAA23354.1; JOINED.
 DR EMBL: AB005524; BAA23354.1; JOINED.
 DR EMBL: AB005525; BAA23354.1; JOINED.
 DR EMBL: X90563; CAA62152.1; ALT_INIT.
 DR EMBL: X90563; CAA62153.1; -;
 DR PIR: S42489; S42489.
 DR PDB: 1FM6; 16-FEB-01.
 DR PDB: 1FM9; 16-FEB-01.
 DR TRANSFAC: T03731; -;
 DR Genew: HGNC:9236; PARG.
 DR MIM: 601487; -;
 DR MIM: 604367; -;
 DR InterPro: IPR000536; Hormone_rec_1lg.
 DR InterPro: IPR001723; Stchrnm_receptor.
 DR InterPro: IPR001628; ZnI_C4steroid.
 DR Pfam: PF00104; hormone_rec; 1.
 DR Pfam: PF00105; zf-C4; 1.
 DR PRINTS: PR00398; STRODHORMONER.
 DR PRINTS: PR00047; STROIDPFINGER.
 DR Prodom: PD000035; ZnI_C4steroid; 1.
 DR SMART: SM00430; HOI1; 1.
 DR SMART: SM00399; ZnF_C4; 1.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor: Transcription regulation; Activator; DNA-binding;
 KW Nuclear protein; Zinc-finger; Multigene family; Alternative splicing;
 KW Phosphorylation; Polymorphism; Disease mutation; Diabetes mellitus;
 KW Obesity; 3D-structure.
 FT DNA_BIND 139 203
 FT ZN_FLNG 139 159
 FT ZN_FLNG 176 198
 FT DOMAIN 318 505
 FT MOD_RES 112 112
 FT VARSPPLIC 1 30
 FT VARIANT 12 12
 FT VARIANT 113 113
 FT VARIANT 314 314
 FT VARIANT 316 316
 FT VARIANT 318 318
 FT VARIANT 495 495
 FT CONFLICT 36 37
 FT CONFLICT 213 214
 FT CONFLICT 240 240
 FT CONFLICT 424 426
 SO SEQUENCE 505 AA; 57620 MW; 3933EF36A0ECAAF CRC64;

Query Match 63.6%; Score 2231; DB 1; Length 505;
 Best Local Similarity 87.7%; Pred. No. 1.7e-116;
 Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 1 MGTTLDSPIDSDSFTDTLSANISQEMTVDTEMPPTNFGISSVDLSWEDHSF 60
 DB 1 MGTTLDSPIDSDSFTDTLSANISQEMTVDTEMPPTNFGISSVDLSWEDHSF 60
 QY 61 DLRPTTVDFSSISTPHYEDIPFTPTDPAVDKYDKLQEOSAIKVPASPPYSEKT 120
 DB 61 DLRPTTVDFSSISTPHYEDIPFTPTDPAVDKYDKLQEOSAIKVPASPPYSEKT 120
 QY 121 QLYN----- 124
 DB 121 QLYN----- 124
 QY 125 -----RNKCQYCRFOCKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDLNPESADLR 178
 DB 181 RJHKSRRNCQYCRFOCKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDLNPESADLR 240
 QY 179 ALAKHLVDSYISFPLTKAKARALLGKTTDKSPFYIYDMNSLMGEGDKIRKHTHPLOE 238
 DB 241 ALAKHLVDSYISFPLTKAKARALLGKTTDKSPFYIYDMNSLMGEGDKIRKHTHPLOE 300
 QY 239 QSKVAIRIFOGQCFRSEAVQETFEYAKSIPGFVNLNDQVTLTKYGVHEIITMLAS 298
 DB 301 QSKVAIRIFOGQCFRSEAVQETFEYAKSIPGFVNLNDQVTLTKYGVHEIITMLAS 360
 QY 299 LMKKGVLSISGQGFMTREFLSLKRPGDEMEKPEFEAVKFNALDSDSLAIFIAVII 358
 DB 361 LMKKGVLSISGQGFMTREFLSLKRPGDEMEKPEFEAVKFNALDSDSLAIFIAVII 420
 QY 359 LSGDRGLNVPRFIEDIONLQALELOLKLNPRESSOLFALLOKMTDLROIVHEVOL 418
 DB 421 LSGDRGLNVPRFIEDIONLQALELOLKLNPRESSOLFALLOKMTDLROIVHEVOL 480
 QY 419 LOVTKETDMSLHPLLOEIKRDLY 443
 DB 481 LOVTKETDMSLHPLLOEIKRDLY 505

RESULT 2
 PPRAT_MACMU STANDARD; PRT; 505 AA.
 ID PPRAT_MACMU STANDARD; PRT; 505 AA.
 AC 018924; 09TOM6;
 DT 16-OCT-2001 (rel. 40, Created)
 DT 16-OCT-2001 (rel. 40, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).
 GN PPARC OR NR1C3.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=adipose tissue;
 RX MEDLINE=99021153; PubMed=9806316;
 RA Hotta K., Gustafson T.A., Yoshitaka S., Ortmeier H.K., Bocklin N.L.,
 Hansen B.C.;
 RT "Relationships of PPARgamma and PPARgamma2 mRNA levels to obesity,
 diabetes and hyperinsulinaemia in rhesus monkeys";
 RL Int. J. Obes. Relat. Metab. Disord. 22:1000-1010(1998).
 CC -!- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
 HYPOGLYCEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
 THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
 OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
 PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
 OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.
 CC -!- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
 CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 AND 2 (SHOWN HERE);
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADIPOSE TISSUE. LOWER IN
 CC LIVER, HEART, KIDNEY, STOMACH, DUODENUM AND COLON.
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NRI SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF033103; AAB87480.1; -;
 CC DR EMBL: AF033343; AAB87482.1; -;
 CC DR EMBL: AF033342; AAB87481.1; -;
 CC DR HSSP: P37231; 1FM9.
 CC DR InterPro: IPR000536; Hormone_rec.1lg.
 CC DR InterPro: IPR001723; sthrrm_receptor.
 CC DR InterPro: IPR001628; znf_C4steroid.
 CC DR Pfam: PF00104; hormone_rec.1.
 CC DR Pfam: PF00105; zf-C4; 1.
 CC DR PRINTS: PR00398; STRDHOMNER.
 CC DR PRINTS: PR00047; STROIDINGER.
 CC DR Prodom: PD000035; znf_C4steroid; 1.
 CC DR SMART: SM00430; HOLY.1.
 CC DR SMART: SM00399; znf_C4; 1.
 CC DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 CC DR Receptor; Transcription regulation; Activator; DNA-binding;
 CC KW Nuclear protein; Zinc-finger; Multigene family; Alternating splicing;
 CC KW Phosphorylation.
 CC FT DNA_BIND 139 203 NUCLEAR RECEPTOR-TYPE.
 CC FT ZN_FING 139 159 C4-TYPE.
 CC FT ZN_FING 176 198 C4-TYPE.
 CC FT DOMAIN 318 505 LIGAND-BINDING (POTENTIAL).
 CC FT MOD_RES 112 112 PHOSPHORYLATION (BY MAPK) (BY
 CC FT SIMILARITY).
 CC FT VARSPLIC 1 30 MISSING (IN ISOFORM 1).
 CC SQ SEQUENCE 505 AA; 57590 MW; 41836a624AAAF942 CRC64;

Query Match 63.2%; Score 2216; DB 1; Length 505;
 Best Local Similarity 86.9%; Pred. No. 1.1e-115;
 Matches 439; Conservative 3; Mismatches 1; Indels 62; Gaps 1;

QY 1 MGTTLDSPIDSDSFTDTLSANISQEMTVDTEMPPTNFGISSVDLSWEDHSF 60
 DB 1 MGTTLDSPIDSDSFTDTLSANISQEMTVDTEMPPTNFGISSVDLSWEDHSF 60
 QY 61 DLRPTTVDFSSISTPHYEDIPFTPTDPAVDKYDKLQEOSAIKVPASPPYSEKT 120
 DB 61 DLRPTTVDFSSISTPHYEDIPFTPTDPAVDKYDKLQEOSAIKVPASPPYSEKT 120
 QY 121 QLYN----- 124
 DB 121 QLYN----- 124
 QY 125 -----RNKCQYCRFOCKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDLNPESADLR 178
 DB 181 RJHKSRRNCQYCRFOCKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDLNPESADLR 240
 QY 179 ALAKHLVDSYISFPLTKAKARALLGKTTDKSPFYIYDMNSLMGEGDKIRKHTHPLOE 238
 DB 241 ALAKHLVDSYISFPLTKAKARALLGKTTDKSPFYIYDMNSLMGEGDKIRKHTHPLOE 300
 QY 239 QSKVAIRIFOGQCFRSEAVQETFEYAKSIPGFVNLNDQVTLTKYGVHEIITMLAS 298
 DB 301 QSKVAIRIFOGQCFRSEAVQETFEYAKSIPGFVNLNDQVTLTKYGVHEIITMLAS 360
 QY 299 LMKKGVLSISGQGFMTREFLSLKRPGDEMEKPEFEAVKFNALDSDSLAIFIAVII 358
 DB 361 LMKKGVLSISGQGFMTREFLSLKRPGDEMEKPEFEAVKFNALDSDSLAIFIAVII 420

QY 359 LSGDRPGLNVKRIEDIONNLQALELQKLNHPRESSQLEFALKLOKMTDLROIVTEHVOL 418
 DB 421 LSGDRPGLNVKRIEDIONNLQALELQKLNHPRESSQLEFALKLOKMTDLROIVTEHVOL 480
 QY 419 LOYIKKETDMSLHPLLOEITYKDL 443
 DB 481 LOYIKKETDMSLHPLLOEITYKDL 505
 RESULT 3
 PRTAT_PIG STANDARD: PRT; 504 AA.
 AC 062807; 077815;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).
 GN PPARG OR NR1C3.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC TISSUE=adipose tissue;
 RX MEDLINE=99132654; PubMed=991452;
 RA Houseknecht K.L., Bidwell C.A., Portocarrero C.P., Spurrlock M.E.;
 RT "Expression and cDNA cloning of porcine peroxisome proliferator-
 RT activated receptor gamma (PPARgamma).";
 RL Gene 225:89-96(1998).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=99410873; PubMed=10481259;
 RA Ding S.T., McNeil R.L., Mersmann H.J.;
 RT "Expression of porcine adipocyte transcripts: tissue distribution and
 RT differentiation in vitro and in vivo.";
 RL Comp. Biochem. Physiol. 123B:307-318(1999).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RC STRAIN=Duroc, and Norwegian Landrace; TISSUE=Adipocyte;
 RX MEDLINE=98401156; PubMed=9731203;
 RA Grindler E., Sundvold H., Klungland H., Lien S.;
 RT "Characterisation of porcine peroxisome proliferator-activated
 RT receptors gamma 1 and gamma 2: detection of breed and age differences
 RT in gene expression.";
 RL Biochem. Biophys. Res. Commun. 249:713-718(1998).
 RL Biochem. Biophys. Res. Commun. 249:713-718(1998).
 CC -!- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
 CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
 CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
 CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
 CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
 CC OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.
 CC -!- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 AND 2 (SHOWN HERE); ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADIPOSE TISSUE AND LOWER
 CC IN SPLEEN. VERY LOW LEVELS IN KIDNEY, INTESTINE, LUNG AND MUSCLE.
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NRI SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF059245; AAC14348.1; -;
 CC EMBL; AF103946; AAD19577.1; -;
 CC EMBL; AJ006757; CAA07225.1; -;

DR EMBL; AJ006756; CAA07224.1; -;
 DR HSSPE; P37231; 1FM9.
 DR Interpro: IPR000536; Hormone_rec.1lg.
 DR Interpro: IPR001723; Stghrm_receptor.
 DR Interpro: IPR001628; ZnfC4steroid.
 DR Pfam: PF00104; hormone_rec.1.
 DR Pfam: PF00105; zf-C4.1.
 DR PRINTS: PR00398; STRDHORMONER.
 DR PRINTS: PR00047; STROIDFINGER.
 DR Prodom: PD000035; ZnfC4steroid.1.
 DR SMART: SM00430; HOL1.1.
 DR SMART: SM00399; ZnfC4.1.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR.1.
 DR Receptor; Transcription regulation; Activator; DNA-binding;
 KW Nuclear protein; zinc-finger; Multigene family; Alternative splicing;
 KW Phosphorylation.
 FT DNA_BIND 138 202 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 138 158 C4-TYPE.
 FT ZN_FING 175 197 C4-TYPE.
 FT DOMAIN 317 504 LIGAND-BINDING (POTENTIAL).
 FT MOD_RES 111 111 PHOSPHORYLATION (BY MAPK) (BY
 FT VARSPLIT 1 29 MISSING (IN ISOFORM 1).
 FT CONFLICT 381 381 K -> R (IN REF. 3).
 FT CONFLICT 426 426 G -> R (IN REF. 3).
 SQ SEQUENCE 504 AA; 57512 MW; 57C0BDE901493D31 CRC64;
 Query Match 61.9%; Score 2171.5; DB 1; Length 504;
 Best Local Similarity 85.5%; Pred. No. 3.2e-113;
 Matches 432; Conservative 6; Mismatches 4; Indels 63; Gaps 2;
 QY 1 MGETLSDSPIDPSDSFTTLTANISQEMTNVDTMPFWPTNGISSVLSVMEHSHSF 60
 DB 1 MGETLSDSLIDPSDAF-DTLNANISQEVMTWDEMPFPTNGISSVLSVWDHSHSF 59
 QY 61 DIKFTTVPDSISTPHYEDIPTRPDVADYKYDLQIYOSAIKVPASPPYSEKT 120
 DB 60 DIKFTTVPDSISTPHYEDIPTRPADPVADYKYDLQIYOSAIKVPASPPYSEKT 119
 QY 121 QLVN----- 124
 DB 120 QLVNKRHEEPSNLSAIECRVCGDKASGFHYGNHACGCKGFFRRTRILKLIYRCDLNC 179
 QY 125 -----RNRKCOYRFQKCLAVGMSHNAIRFGMPQAEKREKLAELISSDIDOLNPESADLR 178
 DB 180 RIHKSRNKCQYCRFQKCLAVGMSHNAIRFGMPQAEKREKLAELISSDIDOLNPESADLR 239
 QY 179 ALAKHLVDSTYISFPLTKAKARAILTGKTTDKSPVITYYMNLSLMDGEDIKRKHITPLOE 238
 DB 240 ALAKHLVDSTYISFPLTKAKARAILTGKTTDKSPVITYYMNLSLMDGEDIKRKHITPLOE 299
 QY 239 OSKEVAIRIFQGGQFVSVEAVOEITFYAKSIGFVNLINDOYTLKYGVHEIITYMLAS 298
 DB 300 OSKEVAIRIFQGGQFVSVEAVOEITFYAKNIGFVNLINDOYTLKYGVHEIITYMLAS 359
 QY 299 LMKKDVLLISEGGFMTREFLSLRKPFQDFMEPKREFAVKNALELDSDLAIFAVIY 358
 DB 360 LMKKDVLLISEGGFMTREFLSLRKPFQDFMEPKREFAVKNALELDSDLAIFAVIY 419
 QY 419 LOYIKKETDMSLHPLLOEITYKDL 443
 DB 480 LOYIKKETDMSLHPLLOEITYKDL 504
 RESULT 4
 PRTAT_MOUSE STANDARD: PRT; 505 AA.
 ID PRTAT_MOUSE
 AC P37238;
 DT 01-OCT-1994 (Rel. 30; Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID:10090;
 RN [1]
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RC TISSUE=adipose tissue;
 RX MEDLINE=95011536; PubMed=7926726;
 RA Tontonoz P., Hu E., Graves R.A., Budavari A.I., Spiegelman B.M.;
 RT "PPAR gamma 2: tissue-specific regulator of an adipocyte enhancer.";
 RL Gene Dev. 8:1224-1234(1994).
 RN [2]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=BALB/c; TISSUE=Heart;
 RX MEDLINE=94059089; PubMed=8240342;
 RA Chen F., Law S.W., O'Malley B.W.;
 RT "Identification of two mppar related receptors and evidence for the
 RT existence of five subfamily members.";
 RL Biochem. Biophys. Res. Commun. 196:671-677(1993).
 RN [3]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;
 RX MEDLINE=94086482; PubMed=8262913;
 RA Zhu Y., Alvares K., Huang Q., Rao M.S., Reddy J.K.;
 RT "Cloning of a new member of the peroxisome proliferator-activated
 RT receptor gene family from mouse liver.";
 RL J. Biol. Chem. 268:26817-26820(1993).
 RN [4]
 RN SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Liver;
 RX MEDLINE=94316694; PubMed=8041794;
 RA Klierer S.A., Forman B.M., Blumberg B., Ong E.S., Borgmeyer U.,
 RA Mangelsdorf D.J., Umesono K., Evans R.M.;
 RT "Differential expression and activation of a family of murine
 RT peroxisome proliferator-activated receptors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:7355-7359(1994).
 RN [5]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=9629427; PubMed=8647948;
 RA Vidal-Puig A., Jimenez-Linan M., Lowell B.B., Hamann A., Hu E.,
 RA Spiegelman B., Flier J.S., Moller D.E.;
 RT "Regulation of PPAR gamma gene expression by nutrition and obesity in
 RT rodents.";
 RL J. Clin. Invest. 97:2553-2561(1996).
 CC -i- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
 CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
 CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
 CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
 CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
 CC OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.
 CC -i- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
 CC -i- SUBCELLULAR LOCATION: Nuclear.
 CC -i- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 AND 2 (SHOWN HERE); ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -i- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADIPOSE TISSUE. ALSO
 CC FOUND IN LIVER, SKELETAL MUSCLE, HEART, ADRENAL GLAND, SPLEEN,
 CC KIDNEY AND INTESTINE.
 CC -i- DEVELOPMENTAL STAGE: IT APPEARS FIRST AT DAY 13.5 POSTCONCEPTION,
 CC AND INCREASES UNTIL BIRTH.
 CC -i- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NRI SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@sib-sib.ch).

CC -----
 DR EMBL: U09138; AAA62277.1;
 DR EMBL: U01664; AAA62110.1;
 DR EMBL: U01841; AAC52134.1;
 DR EMBL: U10374; AAA19971.1;
 DR PIR: JN0881; JN0881.
 DR HSSP: P37231; 1FM9.
 DR TRANSFAC: T02529;
 DR MGD: MGI:97747; Pparg.
 DR InterPro: IPR000536; Hormone_rec_1lg.
 DR InterPro: IPR001723; Stohrm_recceptor.
 DR InterPro: IPR001628; ZnF_C4steroid.
 DR Pfam: PF00104; hormone_rec_1.
 DR Pfam: PF00105; zf-C4_1.
 DR PRINTS: PR00398; STRDHORMONER.
 DR PRINTS: PR00047; STROIDFINGER.
 DR PRODOM: PD000035; ZnF_C4steroid; 1.
 DR SMART: SM00430; HOL1; 1.
 DR SMART: SM00399; ZnF_C4; 1.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 KW Receptor; Transcription regulation; Activator; DNA-binding;
 KW Nuclear protein; zinc-finger; Multigene family; Alternative splicing;
 KW Phosphorylation.
 KW DNA_BIND 139 203 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 139 159 C4-TYPE.
 FT ZN_FING 176 198 C4-TYPE.
 FT DOMAIN 318 505 LIGAND-BINDING (POTENTIAL).
 FT MOD_RES 112 112 PHOSPHORYLATION (BY MAPK) (BY
 FT SIMILARITY).
 FT VARSPLIC 1 30 MISSING (IN ISOFORM 1).
 FT CONFLICT 213 214 MP -> DR (IN REF. 2).
 FT CONFLICT 281 283 NSL -> SSF (IN REF. 2).
 FT CONFLICT 383 383 N -> S (IN REF. 2 AND 4).
 FT CONFLICT 497 497 L -> F (IN REF. 2).
 SQ SEQUENCE 505 AA; 57598 MW; AB8F3F6086E2A10A CRC64;
 Query Match 61.4%; Score 2153; DB 1; Length 505;
 Best Local Similarity 84.0%; Pred. No. 3.4e-112;
 Matches 424; Conservative 9; Mismatches 10; Indels 62; Gaps 1;
 QY 1 MGTTLDSPTIDESDFTTSLNSIQEMTWDTMPNPTNFGISSVDSLWEDHSSEF 60
 DB 1 METTLDSDVDEHGAFAALPMSTSQETIMVDTMPNPTNFGISSVDSLWEDHSSEF 60
 QY 61 DLRPTTVFSSISPTHYDIPPTRPVADYKYDKLOEYOSAKVBPAPPYSEXT 120
 DB 61 DLRPTTVFSSISAPHYEDIPPTRADPVADYKYDKLOEYOSAKVBPAPPYSEXT 120
 QY 121 QLYN----- 124
 DB 121 QLYNRPHEPNSLMAIECRVCGDKASGFHYVHACGEGKFFRTIRLKLIRCDLNC 180
 QY 125 -----RMKCYCRFOKCLAVGSHNAIRFGMPAEEKKLAELISSDIDOLNPESADLR 178
 DB 181 RIHKSRNKCQYCRFOKCLAVGSHNAIRFGMPAEEKKLAELISSDIDOLNPESADLR 240
 QY 179 ALAKHLDYISFPLTKKARAILTGKTTDSPFIYMNLSIMMGEDIKRKHTIPLOE 238
 DB 241 ALAKHLDYISFPLTKKARAILTGKTTDSPFIYMNLSIMMGEDIKRKHTIPLOE 300
 QY 239 OSKEVAIRIFOGCOPRSVAVOEITEYAKSIGFVNLNDQVTLTKGVHEITVMTLAS 298
 DB 301 OSKEVAIRIFOGCOPRSVAVOEITEYAKSIGFVNLNDQVTLTKGVHEITVMTLAS 360
 QY 299 LNNKGVLLISEQGFTRFELSKRPFGEDEPPKFEFVKNALDELDDSLAIFAVYII 358
 DB 361 LNNKGVLLISEQGFTRFELSKRPFGEDEPPKFEFVKNALDELDDSLAIFAVYII 420
 QY 359 LSGDRGLNVAPRIEDIONLQALELQKLNHPSSQLFAKLQKMDLROIVTEHYDL 418
 DB 421 LSGDRGLNVAPRIEDIONLQALELQKLNHPSSQLFAKLQKMDLROIVTEHYDL 480
 QY 419 LQVIRKTTETDMSLHPLLOEIKYKDL 443


```

Db 121 OLYNRPHEPNSLMAIECRVCGDKASGFHYGVHACECGKGFRTIRLKLITYRCDLNC 180
QY 125 -----NRKCOYCRFOCKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDLNPESADLR 178
Db 181 RIHKSRNKCQYCRFOCKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDLNPESADLR 240
QY 179 ALAKHLVDSYIKSFPLTYAKARAILTGKTTDKSPFVIYDMNSLMGDEKIKFKHTPLQE 238
Db 241 ALAKHLVDSYIKSFPLTYAKARAILTGKTTDKSPFVIYDMNSLMGDEKIKFKHTPLQE 300
QY 239 OSKEVAIRIFOGCQFRSVEAVOEITEYAKSIPGFVNLDLNDQVTLTKGVHEIITMLAS 298
Db 301 OSKEVAIRIFOGCQFRSVEAVOEITEYAKNIPGFNLNDLNDQVTLTKGVHEIITMLAS 360
QY 299 LMKNDGVLISEGQCFMTEFLSLRKPFGDEMPKFEFAVKNALDELDDSLAIFIAVII 358
Db 361 LMKNDGVLISEGQCFMTEFLSLRKPFGDEMPKFEFAVKNALDELDDSLAIFIAVII 420
QY 359 LSGDRPGLLNKPRIEDIDNLQALELQKLNHPESSQLFAKLLQKMTDLRQIYTEHVQL 418
Db 421 LSGDRPGLLNKPRIEDIDNLQALELQKLNHPESSQLFAKLLQKMTDLRQIYTEHVQL 480
QY 419 LQVIKKTETDMSLHPLLOEIKDLY 443
Db 481 LQVIKKTETDMSLHPLLOEIKDLY 505

```

RESULT 6

```

PPAT_BOVIN
ID PPAT_BOVIN STANDARD; PRT; 505 AA.
AC 018971;

```

```

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).
GN PPARG OR NR1C3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;

```

```

RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RT TISSUE=Fat;

```

```

RX MEDLINE=98042483; PubMed=9367859;

```

```

RA Sundvold H., Brzozowska A., Lien S.;

```

```

RT "Characterisation of bovine peroxisome proliferator-activated
RT receptors gamma 1 and gamma 2: genetic mapping and differential
RT expression of the two isoforms.";

```

```

RL Biochem. Biophys. Res. Commun. 239:857-861(1997).

```

```

CC -1- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
CC PROXIMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
CC OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.

```

```

CC -1- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.

```

```

CC -1- SUBCELLULAR LOCATION: Nuclear.

```

```

CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 AND 2 (SHOWN HERE);

```

```

CC ARE PRODUCED BY ALTERNATIVE SPLICING.

```

```

CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADIPOSE TISSUE. LOWER IN
CC SPLEEN AND LUNG. ALSO DETECTED IN OVARY MAMMARY GLAND AND SMALL
CC INTESTINE.

```

```

CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

```

```

CC NRI SUBFAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: Y12420; CAAT3033.1;
DR EMBL: Y12419; CAAT3032.1;
DR HSSP: P37231; 1FM9.
DR TRANSFAC: T04780;
DR InterPro: IPR000536; Hormone_rec_1lg.
DR InterPro: IPR001723; Stchrnm_receptor.
DR InterPro: IPR001628; znf_C4steroid.
DR Pfam: PF00104; hormone_rec.1.
DR Pfam: PF00105; znf_C4.1.
DR PRINTS: PR00398; STRDHORMONER.
DR PRINTS: PR00047; STRIDFINGER.
DR ProDom: PD000035; znf_C4steroid.1.
DR SMART: SM00430; HOLT.1.
DR SMART: SM00399; znf_C4.1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; Activator; DNA-binding;
KW Nuclear protein; Zinc-finger; Multigene family; Alternative splicing;
KW Phosphorylation.
FT DNA_BIND 139 203 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 139 159 C4-TYPE.
FT ZN_FING 176 198 C4-TYPE.
FT DOMAIN 318 505 LIGAND-BINDING (POTENTIAL).
FT MOD_RES 112 112 PHOSPHORYLATION (BY MAPK) (BY
FT SIMILARITY).
FT VARSPLIC 1 30 MISSING (IN ISOFORM 1).
FT SEQUENCE 505 AA; 57579 MW; 5F20B115087B3C63 CRC64;

```

```

Query Match 60.8%; Score 2132; DB 1; Length 505;
Best Local Similarity 83.6%; Pred. No. 4,9e-111;
Matches 422; Conservative 11; Mismatches 10; Indels 62; Gaps 1;

```

```

QY 1 MGFTLGDSPIDPESDSEFDFLSANISOEMTVDTEMPTNFGISSVDLSMDEHSHSF 60
Db 1 MGFTLGDALDPSESPPEFVTSARTSQETVMDTEMPWPNFGISSVDLSMDDHSHAF 60
QY 61 DIKPFYTVDFSSISTPHXEDIPFTRTPDVADVADKYDLKLOEYQSAIKYEPASPPYSEKT 120
Db 61 DIKPFYTVDFSSISTPHXEDIPFTRTPDVADVADKYDLKLOEYQSAIKYEPSPYISSEKT 120
QY 121 QLY----- 123
Db 121 QLYSPHEPNSLMAIECRVCGDKASGFHYGVHACECGKGFRTIRLKLITYRCDLNC 180
QY 124 -----NRKCOYCRFOCKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDLNPESADLR 178
Db 181 RIHKSRNKCQYCRFOCKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDLNPESADLR 240
QY 179 ALAKHLVDSYIKSFPLTYAKARAILTGKTTDKSPFVIYDMNSLMGDEKIKFKHTPLQE 238
Db 241 ALAKHLVDSYIKSFPLTYAKARAILTGKTTDKSPFVIYDMNSLMGDEKIKFKHTPLQE 300
QY 239 OSKEVAIRIFOGCQFRSVEAVOEITEYAKSIPGFVNLDLNDQVTLTKGVHEIITMLAS 298
Db 301 OSKEVAIRIFOGCQFRSVEAVOEITEYAKNIPGFNLNDLNDQVTLTKGVHEIITMLAS 360
QY 299 LMKNDGVLISEGQCFMTEFLSLRKPFGDEMPKFEFAVKNALDELDDSLAIFIAVII 358
Db 361 LMKNDGVLISEGQCFMTEFLSLRKPFGDEMPKFEFAVKNALDELDDSLAIFIAVII 420
QY 359 LSGDRPGLLNKPRIEDIDNLQALELQKLNHPESSQLFAKLLQKMTDLRQIYTEHVQL 418
Db 421 LSGDRPGLLNKPRIEDIDNLQALELQKLNHPESSQLFAKLLQKMTDLRQIYTEHVQL 480
QY 419 LQVIKKTETDMSLHPLLOEIKDLY 443
Db 481 LQVIKKTETDMSLHPLLOEIKDLY 505

```

```

RESULT 7
PPAT_CRIGR
ID PPAT_CRIGR STANDARD; PRT; 475 AA.
AC P57797;

```

```

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).
GN PPARC OR NR1C3.
OS Citellus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Citellus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Liver;
RX MEDLINE=96032361; PubMed=7557447;
RA Apelo C., Poonnec P., Saladin R., Auwerx J., Boulukos K.E.;
RT "cDNA cloning and characterization of the transcriptional activities
RT of the hamster peroxisome proliferator-activated receptor hppar
RT gamma."
RL Gene 162:297-302(1995).
CC -1- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
CC OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.
CC -1- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NRI SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z30972; CAA83219.1; -.
DR HSSP: P37231; 1FM9.
DR InterPro: IPR000536; Hormone_rec.1lg.
DR InterPro: IPR001723; Stchrnm_receptor.
DR InterPro: IPR001628; Znfc4steroid.
DR Pfam: PF00104; hormone_rec.1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR00398; STRDHOMONER.
DR PRINTS: PR00047; STRDIDFINGER.
DR PRODOM: PR000035; Znfc4steroid.1.
DR SMART: SM00430; HOL1; 1.
DR SMART: SM00399; Znfc4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; Activator; DNA-binding;
KW Nuclear protein; zinc-finger; Multigene family; Phosphorylation.
FT DNA_BIND 109 173 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 109 129 C4-TYPE.
FT ZN_FING 146 168 C4-TYPE.
FT DOMAIN 288 475 LIGAND-BINDING (POTENTIAL).
FT MOD_RES 82 82 PHOSPHORYLATION (BY MAPK) (BY
FT SIMILARITY)
SO SEQUENCE 475 AA; 54472 MW; BBDCA0704F837ADB CRC64;

Query Match 58.2%; Score 2043; DB 1; Length 475;
Best Local Similarity 84.8%; Pred. No. 3.7e-106;
Matches 403; Conservative 6; Mismatches 4; Indels 62; Gaps 1;

```

```

QY 125 -----RNCQYCRFOKCLAVGMSHNAIRF 148
DB 121 YGVNAHEGCGKGFRRRTIRKLTYDRCDLNCRIHKSRKQCRCFROKCLAVGMSHNAIRF 180
QY 149 GMPQAEKEKLLAEISSDIDQNPESADLRALAKLYSYIKSPFLTKAKAAITLGT 208
DB 181 GMPQAEKEKLLAEISSDIDQNPESADLRALAKLYSYIKSPFLTKAKAAITLGT 240
QY 209 DKSPFVIYDMSLMNGEDKIKFKHTTLPLOESKEVAIRIFQCGCPRSYEAQVEITEYAKS 268
DB 241 DKSPFVIYDMSLMNGEDKIKFKHTTLPLOESKEVAIRIFQCGCPRSYEAQVEITEYAKN 300
QY 269 IPGFVNLNDQVTLTKYGVHEIIVTMLASLMNKDGVLSGCGFMTREPLKSLRPPGD 328
DB 301 IPGFVNLNDQVTLTKYGVHEIIVTMLASLMNKDGVLSGCGFMTREPLKSLRPPGD 360
QY 329 FMEPRFEFAVKFNALIEDSDLAIFAVIILSGDRPGLLNKPRIDIDNLLQALELQK 388
DB 361 FMEPRFEFAVKFNALIEDSDLAIFAVIILSGDRPGLLNKPRIDIDNLLQALELQK 420
QY 389 LNHPESSQLFAKLQKMTDLROIVTEHVQLLOVYIKTETDMSLHPLQEIYKDY 443
DB 421 LNHPESSQLFAKLQKMTDLROIVTEHVQLLOVYIKTETDMSLHPLQEIYKDY 475

RESULT 8
PSTAT RABIT
ID PSTAT RABIT STANDARD: PRT: 475 AA.
AC 019052;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).
GN PPARC OR NR1C3.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white;
RX MEDLINE=97418978; PubMed=9275054;
RA Michael L.F., Lazar M.A., Mendelson C.R.;
RT "Peroxisome proliferator-activated receptor gamma expression is
RT induced during cyclic adenosine monophosphate-stimulated
RT differentiation of alveolar type II pneumocytes."
RT Endocrinology 138:3695-3703(1997).
CC -1- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
CC OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.
CC -1- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NRI SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U84893; AAB96380.1; -.
DR HSSP: P37231; 1FM9.
DR InterPro: IPR000536; Hormone_rec.1lg.
DR InterPro: IPR001723; Stchrnm_receptor.
DR InterPro: IPR001628; Znfc4steroid.
DR Pfam: PF00104; hormone_rec.1.
DR Pfam: PF00105; zf-C4; 1.

```



```
|||||
Db 180 NAIRFRMPAEKEKLEIAISSDIDLPESADQRYLAKHLYSYKSFPLTKAKAGHP 239
Cc 204 TGT--TDSPEFYIYMNLSLMGEDIKFKHPILOEO-----SKEVAIRIFGCGCF 253
Cc 240 DQSHRONSGYRHRHADDGGSDGAVRE--PRAEQGGSDNLPALVALR----- 290
Cc 254 RSEVAOETIEYAKSIPGFVNLNDQVTLTKYGVHEIITMLASLMKNGVLISEGGF 313
Cc 291 ---GVRTEIEFAKNIPGFVSLDNLQVTLTKYGVHEIITMLASLMKNGVLIAGEGCF 347
Cc 314 MTEFLKSLKRPDMEKPEFPAVKFNALEDDSLAIFIAVIIISGDRPGLLNKPIE 373
Cc 348 MTEFLKSLKRPDMEKPEFPAVKFNALEDDSLAIFIAVIIISGDRPGLLNKPIE 407
Cc 374 DIDNLLQALELQKLNHPSSQLPAKLLQKMDLQIYVEHQVLQVTKETEDMSLP 433
Cc 408 DIDSLQALELQKLNHPSSQLPAKLLQKMDLQIYVEHQVLQVTKETEDMSLP 467
Cc 434 LLOEYKDLV 443
Cc 468 LLOEYKDLV 477

RESULT 10
PPAS_HUMAN STANDARD; PRT; 441 AA.
ID 003181;
AC 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peroxisome proliferator activated receptor beta (PPAR-beta)
DE (PPAR-delta) (Nuclear hormone receptor 1) (NUC1) (NUC1).
GN PPARB OR NR1C2 OR PPARD.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE=93078797; PubMed=1333051;
RA Schmidt A., Endo N., Rutledge S.J., Vogel R., Shinar D.,
RA Rodan G.A.;
RT "Identification of a new member of the steroid hormone receptor
RT superfamily that is activated by a peroxisome proliferator and fatty
RT acids."
RL Mol. Endocrinol. 6:1634-1641(1992).
RN 12;
RP SEQUENCE FROM N.A.
RX MEDLINE=20311491; PubMed=10851270;
RA Skogsberg J., Kannisto K., Roshani L., Gagne E., Hamsten A.,
RA Larsson C., Ehrenborg E.;
RT "Characterization of the human peroxisome proliferator activated
RT receptor delta gene and its expression."
RL Int. J. Mol. Med. 6:73-81(2000).
RN 13;
RP SEQUENCE FROM N.A.
RA Phillips S.;
RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
Cc -!- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
Cc HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
Cc THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
Cc OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
Cc PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS.
Cc -!- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
Cc -!- SUBCELLULAR LOCATION: Nuclear.
Cc -!- TISSUE SPECIFICITY: UBICUITOUS WITH MAXIMAL LEVELS IN PLACENTA AND
Cc SKELETAL MUSCLE.
Cc -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
Cc NRI SUBFAMILY.
Cc -----
Cc This SWISS-PROT entry is copyright. It is produced through a collaboration
Cc between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (See http://www.isb-sib.ch/announce/
Cc or send an email to license@isb-sib.ch).
Cc -----
Cc EMBL; L07592; AAA36469.1; -.
Cc EMBL; AF246303; AAF62553.1; -.
Cc EMBL; AF246299; AAF62553.1; JOINED.
Cc EMBL; AF246300; AAF62553.1; JOINED.
Cc EMBL; AF246301; AAF62553.1; JOINED.
Cc EMBL; AF246302; AAF62553.1; JOINED.
Cc EMBL; AL022721; CAB38629.1; -.
Cc PIR; A45360; A45360.
Cc HSSP; P03372; LHCO.
Cc TRANSFAC; T02745; -.
Cc Genew; HGNC:9235; PPARD.
Cc MIM; 600409; -.
Cc InterPro; IPR000536; Hormone_rec_1lg.
Cc InterPro; IPR001723; Stdrhm_receptor.
Cc InterPro; IPR001628; Znf_C4steroid.
Cc Pfam; PF00104; hormone_rec.1.
Cc Pfam; PF00105; zf-C4; 1.
Cc PRINTS; PR00396; STRDHOMONER.
Cc PRODOM; PD000035; Znf_C4steroid.1.
Cc SMART; SM00399; ZNF_C4; 1.
Cc SMART; PS00031; NUCLEAR_RECEPTOR.1.
Cc Receptor; Transcription regulation; Activator; DNA-binding;
Cc Nuclear protein; zinc-finger; Multigene family.
Cc DNA_BIND 74 138 NUCLEAR RECEPTOR-TYPE.
Cc FT ZN_FING 74 94 C4-TYPE.
Cc FT ZN_FING 111 133 C4-TYPE.
Cc FT DOMAIN 254 441 LIGAND-BINDING (BY SIMILARITY).
Cc SQ SEQUENCE 441 AA; 49903 MW; 94FBB2A4B6521E8 CRC64;

Query Match 31.4%; Score 1101.5; DB 1; Length 441;
Best Local Similarity 63.8%; Pred. No. 3.5e-54;
Matches 208; Conservative 56; Mismatches 61; Indels 1; Gaps 1;

Cc 119 KTYLYRNKCYQCFQKCLAVGNSHNAIRGRMPQAEKELLAET--SSDIDLPESADL 177
Cc 116 KTKKRNKRCQYCFQKCLAVGNSHNAIRGRMPQAEKELLAET--SSDIDLPESADL 175
Cc 178 RALAKHLYSYKSPFLTAKARAILTGKTDKSPVYIDMSLMNGGCEKFKHHTPIQ 237
Cc 176 KAFSKHYNAVLEKFWMTKKARSILTGKASTAPRVIHDIETLMOAEGVWKQVNGU 235
Cc 238 EQSKEVAIRIFGCGQFRSEVAOETIEYAKSIPGFVNLNDQVTLTKYGVHEIITMLA 297
Cc 236 PYKEISVHVFRCQCTYETVELTEFPAKSIPFSFLNDQVTLTKYGVHEIITMLA 295
Cc 298 SLNKKGVLISSGQGFMTREFLKSRLKPPGDEMPEKPEFPAVKFNALEDDSLAIFIAVI 357
Cc 296 SYNKGDLVANGSGVFTEFLKSPSDIIEPFEEVAFVNALEDDSLAIFIAI 355
Cc 358 ILSGDRPGLLNKPIEDIDNLLQALELQKLNHPSSQLPAKLLQKMDLQIYVEHQ 417
Cc 356 ILGDRPGLLNKPIEDIDNLLQALELQKLNHPSSQLPAKLLQKMDLQIYVEHQ 415
Cc 418 LLOVIRKTTEDMSLHPLEIYKDLV 443
Cc 416 MNGRIKKTETETSLHPLLEIYKDMY 441

RESULT 11
PPAR_RAT STANDARD; PRT; 468 AA.
ID P37230;
AC 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
```


OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
 PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS.
 CC SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LIVER, KIDNEY AND
 CC HEART. VERY WEAKLY EXPRESSED IN BRAIN AND TESTIS.
 CC -1- DEVELOPMENTAL STAGE: IT APPEARS FIRST AT DAY 13.5 POSTCONCEPTION,
 CC AND INCREASES UNTIL BIRTH.
 CC -1- DISEASE: PEROXISOME PROLIFERATORS ARE A DIVERSE GROUP OF
 CC CHEMICALS THAT INCLUDE HYPOLIPIDEMIC DRUGS, HERBICIDES AND
 CC INDUSTRIAL PLASTICISERS. ADMINISTRATION OF THESE CHEMICALS TO
 CC ROBERTS RESULTS IN THE DRAMATIC PROLIFERATION OF HEPATIC
 CC PEROXISOMES AS WELL AS LIVER HYPERTROPHIA.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NRI SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: X57638; CAA40856.1; -;
 DR EMBL: X75289; CAA53042.1; -;
 DR EMBL: X75290; CAA53042.1; JOINED.
 DR EMBL: X75291; CAA53042.1; JOINED.
 DR EMBL: X75292; CAA53042.1; JOINED.
 DR EMBL: X75293; CAA53042.1; JOINED.
 DR EMBL: X75294; CAA53042.1; JOINED.
 DR EMBL: X89577; CAA61754.1; -;
 DR PIR: S11659; S11659.
 DR HSSP: P03372; IHCO.
 DR TRANSFAC: T00694; -;
 DR MGD: MGI:104740; Ppara.
 DR InterPro: IPR000536; Hormone_rec_11g.
 DR InterPro: IPR001723; Steroid_receptor.
 DR InterPro: IPR001628; znf_C4steroid.
 DR Pfam: PF00104; hormone_rec_1.
 DR Pfam: PF00105; zf-C4_1.
 DR PRINTS: PR00398; STRDHOMONER.
 DR PRINTS: PR0047; STROIDFINGER.
 DR ProDom: PD000035; znf_C4steroid_1.
 DR SMART: SM00430; HOL1; 1.
 DR SMART: SM00399; znf_C4_1.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR_1.
 DR Receptor: Transcription regulation; Activator; DNA-binding;
 KM Nuclear protein; Zinc-finger; Multigene family.
 KM DNA_BIND 102 166 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 102 122 C4-TYPE.
 FT ZN_FING 139 161 C4-TYPE.
 FT DOMAIN 281 468 LIGAND-BINDING (POTENTIAL).
 FT CONFLICT 75 75 A->R (IN REF.1).
 SQ SEQUENCE 468 AA: 52347 MW: 2350A519C610B6B CRC64;

Query Match 30.9%; Score 1083.5; DB 1; Length 468;
 Best Local Similarity 48.4%; Pred. No. 3-8e-53;
 Matches 236; Conservative 65; Mismatches 112; Incls 75; Gaps 9;

QY 7 DSPIDP-----ESDFTDILSANISODMTWDTMPMPPTNFISSVSDYSVMEDHSIEDI 62
 DB 5 ESDIPPLSPLEADDESPLSEFLQEMG-----NIOEISOSIGSESSGSGFG 51
 QY 63 KFE-----TYDPSISPHYEDIPRTDPVADVADKLQGYOSAIKYE 109
 DB 52 ADYQILGSGPGESEGYITDTLSPASSPSVSCPVI---PASTD-----ESPGSLNTE 101
 QY 110 -----PASPPY-----SEKTOLYNNKCOYCRFOKC 136
 DB 102 CRICGDKASGYHYGVHACGCGKGFRRTRILKLYDKCRSKICKKNNKCOYCRFHNC 161
 QY 137 IAVGSHNAIRGRMPQAEKELAEI-SSDIDQINPESADLRALAKLYDSYIKSFPLT 195

DB 162 LSVGSHNAIRGRMPRSEKAKLKAELLCEHDLNDETFADLSGKRHEHYLNFMNN 221
 QY 196 KAKARILFGKTTDSKSPFIYDMNSLMGEDIKFKHTPLQGSKEVAIRFOGCOPRS 255
 DB 222 KVKARVILGAKTSNNPPYIHMETLCNAEKTLYAKMVAANGYE-DKEAEVRFHCCQCKS 280
 QY 256 VEAQVETETAKSIPGFVNLNDQVTLTKYGVHIIYTMLASLNKDKGLISEGQFMT 315
 DB 281 VETVELTFEFAALIGFALNDQVTLTKYGVETLFTMLSSLNKDKGLIANGFIT 340
 QY 316 REFLKSLKRFQEDMEPEKFEFAVKFNALFELDDSLAIFLAVIILSGRGLNVPIDFI 375
 DB 341 REFLNLRKRFQEDIMEPEKFEFAVKFNALFELDDSLAIFLAVIILSGRGLNVPIDFI 400
 QY 376 QNLLQALELQKLNHPSSQLEFAKLQKMDLQRYVMEHOLLOVIRKTTENDSLHPL 435
 DB 401 QEGIVHVLKHLQSNHPDQTFLEPKLQKMDLQRYVMEHOLLOVIRKTTENDSLHPL 460
 QY 436 QEYKDYLY 443
 DB 461 QEYKDYLY 468

RESULT 13
 PPAS_MOUSE STANDARD: PRT: 440 AA.
 ID PPAS_MOUSE P37239;
 AC P35396; P37239;
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-JUN-1994 (Rel. 39, Last annotation update)
 DE 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Peroxisome proliferator activated receptor beta (PPAR-beta)
 DE (PPAR-delta) (Nuclear hormone receptor 1) (NUC1).
 GN PPARB OR NR1C2 OR PPAR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Adipocyte;
 RX MEDLINE=95138211; PubMed=7836471;
 RA Amri E.-Z., Bonhio F., Allaud G., Abumrad N.A., Grimaldi P.A.;
 RT "Cloning of a protein that mediates transcriptional effects of fatty
 RT acids in preadipocytes. Homology to peroxisome proliferator-activated
 RT receptors.";
 RL J. Biol. Chem. 270:2367-2371(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=94316694; PubMed=8041794;
 RA Klierer S.A., Forman B.M., Blumberg B., Ong E.S., Borgmeyer U.,
 RT Mangelsdorf D.J., Unesono K., Evans R.M.;
 RT "Differential expression and activation of a family of murine
 RT peroxisome proliferator-activated receptors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:7355-7359(1994).
 RN [3]
 RP SEQUENCE OF 1-145 FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Brain;
 RX MEDLINE=94059089; PubMed=8240342;
 RA Chen F., Law S.W., O'Malley B.W.;
 RT "Identification of two mPPAR related receptors and evidence for the
 RT existence of five subfamily members.";
 RL Biochem. Biophys. Res. Commun. 196:671-677(1993).
 CC -1- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
 CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
 CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-CoA
 CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
 CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS.
 CC -1- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- TISSUE SPECIFICITY: HEART, ADRENAL AND INTESTINE.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

```

CC      NR1 SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: L28116; AAA6394.1; -.
DR      EMBL: U10375; AAA19972.1; -.
DR      EMBL: U01665; AAA03332.1; ALT_INIT.
DR      HSSP: P03372; IHCO.
DR      TRNSPAC: T04781; -.
DR      MGD: MGI:101884; Ppard.
DR      InterPro: IPR000536; Hormone_rec_1lg.
DR      InterPro: IPR001723; Sterhnm_receptor.
DR      Pfam: PF00104; hormone_rec; 1.
DR      Pfam: PF00105; zf-C4; 1.
DR      PRINTS: PR00398; STROHORMONER.
DR      PRINTS: PR00047; STROIDFINGER.
DR      PRODOM: PD000035; Znf_C4steroid; 1.
DR      SMART: SM00430; HOL1; 1.
DR      SMART: SM00399; Znf_C4; 1.
DR      PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DR      KMW: Receptor; Transcription regulation; Activator; DNA-binding;
KW      Nuclear protein; Zinc-finger; Multigene family.
FT      DNA_BIND 73      137      NUCLEAR RECEPTOR-TYPE.
FT      ZN_FING 73      93      C4-TYPE.
FT      ZN_FING 110     132      C4-TYPE.
FT      DOMAIN 253     440      LIGAND-BINDING (POTENTIAL).
FT      CONFID 149     150      EA -> DG (IN REF. 2).
SQ      SEQUENCE 440 AA; 49715 MW; 58E0F595DD1933DA CRC64;

Query Match      30.9%; Score 1082.5; DB 1; Length 440;
Best Local Similarity 62.6%; Pred. No. 3.9e-53;
Matches 204; Conservative 59; Mismatches 62; Indels 1; Gaps 1;

```

```

GN      PPARA OR NR1C1 OR PPAR.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Liver;
RX      MEDLINE=93277839; PubMed=7684926;
RT      Sher T., Yi H.F., McBride O.W., Gonzales F.J.;
RL      Biochemistry 32:5598-5604(1993).
RN      [2]
RP      SEQUENCE FROM N.A.
RL      Roberts R.A., James N.H., Woodvatt N.J., Macdonald N., Tugwood J.D.;
RN      Submitted (Aug-1996) to the EMBL/Genbank/DBD databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RC      TISSUE=Liver;
RX      MEDLINE=95071923; PubMed=7981125;
RT      Mukherjee R., Jow L., Noonan D., McDonnell D.P.;
RT      "Human and rat peroxisome proliferator activated receptors (PPARs)
RT      demonstrate similar tissue distribution but different responsiveness
RT      to PPAR activators."
RL      J. Steroid Biochem. Mol. Biol. 51:157-166(1994).
RN      [4]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20057165; PubMed=10591208;
RA      Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
RA      Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.R.,
RA      Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasly O.P.,
RA      Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,
RA      Burrill W.D., Burton C., Carter N.P., Chen Y., Clark G.,
RA      Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA      Conroy D., Cordry N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA      Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA      Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA      Gilbert J.G.R., Goward M.E., Graham D.V., Griffiths M.N.D., Hall C.,
RA      Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA      Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA      Lalit G.K., Langford C.F., Leverisha M.A., Lloyd C., Lloyd D.M.,
RA      Marilyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RA      Mcclay J., McLaren S., Mcmurray A.A., Milne S.A., Mortimore B.J.,
RA      Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RA      Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,
RA      Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA      Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RA      Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA      Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA      Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA      Minoshima A., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA      Shintani A., Shibuya K., Yoshitake Y., Aoki N., Mitsuyma S.,
RA      Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA      Dorman A., Fung F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA      Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RA      Phan S., Qi S., Qian Y., Ray L., Ren Q., Shauli S., Sloan D., Song L.,
RA      Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA      Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,
RA      Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA      Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA      Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA      Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Neilson J.,
RA      Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
RA      Emanuel B.S., Shaikh T., Kurahashi H., Saita S., Budarf M.L.,
RA      McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
RA      Kim U.J., Shizuya H., Simon M.I., Dunamski J.P., Peyrard M., Kedra D.,
RA      Serousi E., Franssen I., Tapia I., Brudner C.E., O'Brien K.P.,
RA      Wilkerson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA      Tikhonov Y., Wright H.;
RT      "The DNA sequence of human chromosome 22."
RT      Nature 402:489-495(1999).
CC      -i- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS

```

```
CC HYPO-LIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,  
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA  
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE  
CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS.  
CC -1- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- TISSUE SPECIFICITY: SKELETAL MUSCLE, LIVER, HEART AND KIDNEY.  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
CC NRI SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
CC EMBL: L02932; AAA36468.1; -  
CC EMBL: Y07619; CA68898.1; -  
CC DR EMBL: S74349; AAB32649.1; -  
CC DR EMBL: AL049856; CAB42862.2; -  
CC DR EMBL: AL078611; CAB44427.1; -  
CC DR PIR: A42889; A49289.  
CC HSSP: P03372; IHCO.  
CC TRANSFAC: T02726; -  
CC DR GENE: HGN:9232; PPARA.  
CC MIM: 170998; -  
CC DR InterPro: IPR000536; Hormone_rec_11g.  
CC DR InterPro: IPR001723; Stdhm_receptor.  
CC DR InterPro: IPR001628; Znf_C4steroid.  
CC DR Pfam: PF00104; hormone_rec; 1.  
CC DR Pfam: PF00105; zf-C4; 1.  
CC DR PRINTS: PR00398; STRDHOMONER.  
CC DR PRINTS: PR00047; STROIDFINGER.  
CC DR ProDom: PD000035; Znf_C4steroid; 1.  
CC DR SMART: SM00430; HOL1; 1.  
CC DR SMART: SM00399; Znf_C4; 1.  
CC DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.  
CC DR Receptor: Transcription regulation: Activator: DNA-binding;  
CC KW Nuclear protein; zinc-finger; Multigene family.  
CC KM Nuclear protein; zinc-finger; Multigene family.  
CC FT DNA_BIND 102 166 NUCLEAR RECEPTOR-TYPE.  
CC FT ZN_FING 139 122 C4-TYPE.  
CC FT ZN_FING 102 122 C4-TYPE.  
CC FT DOMAIN 281 468 LIGAND-BINDING (POTENTIAL).  
CC FT CONFLICT 71 71 T -> M (IN REF. 2).  
CC FT CONFLICT 123 123 K -> M (IN REF. 2).  
CC FT CONFLICT 268 268 A -> V (IN REF. 1).  
CC FT CONFLICT 296 296 G -> A (IN REF. 1).  
CC FT CONFLICT 444 444 V -> A (IN REF. 2).  
CC SQ SEQUENCE 468 AA; 52225 MW; 850846FD51ADA883 CRC64;  
  
Query Match 30.7%; Score 1076.5; DB 1; Length 468;  
Best Local Similarity 49.8%; Pred. No. 9; 2e-53;  
Matches 236; Conservative 65; Mismatches 106; Indels 67; Gaps 9;  
  
QY 31 MVTETMPFMTN-----FGISSVDSLVMEDHSHSDIKPFTY---68  
DB 1 MVTETSLCLSLPAGDLESPLSEFLDMGNIQETISGIESGSSFGFTETIYQVSGC 60  
QY 69 ---DFSSIS-TPHYEDIPRTPTDVADYKYLKLOYOS-AIKVE-----PASPPYV-116  
DB 61 PGSDGSVITPTLSPASSPSVTVPPVG-----SVDESPPGALNIEGRICGDKASGYHYG 115  
QY 117 -----SEKTOLYNRNKQYCRPOKCLAVGMSHNAIRFGR 150  
DB 116 VHACEGCKGFRRRTIRKLIVYDKDCRSCKIOKRNRMKQYCRFHKCLVGMSHNAIRFGR 175  
QY 151 MPAQEKELIAEI-SSDIDLPNESAIDRALAHLDYSIKSPFLTKAKARALITGTTD 209  
DB 176 MPSESEKAKLAELITLCHDIESETADLKSLARIYEALKNRMNMKVKARVLISGASN 235  
QY 210 KSPFVIYDMNSLMAGEDKIKFKHITPLQEOSKEVAIRIFGCGQFRSVEAVQETEVAKSI 269
```

```
DB 236 NPPVHIDMETLCAEKTIVAK-LVANGIONKEAEVAFHFCOCTSVETITELFEFKAI 294  
QY 270 PGEVNDLNDQVTLTKGYVHEIITYTLASLMNKDGLVISEGCGFMTRFLSKLRPGDF 329  
DB 295 PGFANLNDQVTLTKGYVEALFAMLSVYMKNDGMVLAVNGNFTREFPLSKLRPGCDI 354  
QY 330 MEPEFEFAVKNLEDDSPALFANVILISGDRPGLNKPREDIODNLQALELQKL 389  
DB 355 MEPEFDPAKFNLEDDSDISLFAVAIICCGDPGLNKGHEKMDGIVHVRHLQOS 414  
QY 390 NHPESSOLFALQKMTDLRQIYTEHVQLQVIRKTEETDMSLPLLOEIKDLY 443  
DB 415 NHPDPIFLPKLLQKNADLQVTEHQVLOIITKTESDAALPHLQLEIRDMY 468  
  
RESULT 15  
PPAR_XENLA  
ID PPAR_XENLA STANDARD; PRT; 474 AA.  
AC P37232;  
DT 01-OCT-1994 (Rel. 30, Last Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Peroxisome proliferator activated receptor alpha (PPAR-alpha).  
GN PPARA OR NR1C1.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;  
OX Xenopodinae; Xenopus.  
NCBI_TaxID=6355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92191267; PubMed=1312391;  
RA Dreyer C., Krey G., Keller H., Givel F., Helltenbein G., Wahl W.;  
RT "Control of the peroxisomal beta-oxidation pathway by a novel family  
RL of nuclear hormone receptors.";  
RN Cell 68:879-887 (1992).  
RP CHARACTERIZATION.  
RX MEDLINE=94100165; PubMed=8274443;  
RA Krey G., Keller H., Mahfoudi A., Medin J., Ozato K., Dreyer C.,  
RA Wahl W.;  
RT "Xenopus peroxisome proliferator activated receptors: genomic  
RT organization, response element recognition, heterodimer formation  
RT with retinoid x receptor and activation by fatty acids.";  
RL J. Steroid Biochem. Mol. Biol. 47:65-73 (1993).  
CC -1- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS  
CC HYPO-LIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,  
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA  
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE  
CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS.  
CC -1- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- TISSUE SPECIFICITY: UBIQUITOUS.  
CC -1- DEVELOPMENTAL STAGE: OCYTES, EMBRYOS, AND ADULTS.  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
CC NRI SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
CC EMBL: M84161; AAA49935.1; -  
CC DR PIR: A42214; -  
CC HSSP: P03372; IHCO.  
CC TRANSFAC: T01352; -  
CC DR InterPro: IPR000536; Hormone_rec_11g.  
CC DR InterPro: IPR001723; Stdhm_receptor.  
CC DR InterPro: IPR001628; Znf_C4steroid.
```

DR	PFam: PF00104, hormone_rec. 1.
DR	Pfam: PF00105, zf-C4. 1.
DR	PRINTS: PR00398, STRDHOMONER.
DR	PRINTS: PD00047, STROIDINER.
DR	ProDom: PD000035, znf_Casteroid. 1.
DR	SMART: SMO0430; HOL1. 1.
DR	SMART: SMO0399; znf_C4. 1.
DR	PROSITE: PS00031, NUCLEAR_RECEPTOR. 1
KW	Receptor; transcription regulation; Activator; DNA-binding;
KW	Nuclear protein; zinc-finger; Multigene family.
FT	DN_BIND 109 173 NUCLEAR RECEPTOR-TYPE.
FT	ZN_FING 109 129 C4-TYPE.
FT	ZN_FING 146 168 C4-TYPE.
FT	DOMAIN 287 474 LIGAND-BINDING (POTENTIAL).
SO	SEQUENCE 474 AA; 5286 MW; 4DBB94131BD43483 CRC64;

Query Match	30.2%	Score 1060.5	DB 1	Length 474
Best Local Similarity	63.5%	Pred. No. 7.1e-52		
Matches 207, Conservative	51	Mismatches 65	Indels 3	Gaps 3

QY	119	KTOLYNNRKCQYCFEQLACWASHNARIKRGMPQAKKELTLEI-SDDIDLNPSADJ	177
Db	151	KIQKNNNRKCQYCFEQLACWASHNARIKRGMRPSKALKLKEVLMCDQVYDSOMADL	210
QY	178	RALAKHLDYSTIKSEFPLTKAKARAILTGKTTDSPEVIYDMSLMMGEDKIKRKHITPLQ	237
Db	211	LSLRLIYDALIKFNMMNKKVKARAILTGKASN--PPEVIHDMETLMAEKLVAK-LVANG	268
QY	238	EQSEKVAIRIRFQOGQFSPSEVAOETTEYKASIFGEVYALDNDQVTLTKYVGEHIIYTLA	297
Db	269	IQNKAEAVRIFHCQCYSVEYVELTTEFAASTFQETFLDNDQVTLTKYVYANFRAMLA	328
QY	298	SLMNKDGVLISGGGFMTRFELSLRRPFGDEMEKPEFAVKFNALDELDDSLAIFAVI	357
Db	329	SVNMKGDMLVYANGFTTRFELSLRRPIDDMEPKPEFMAKFNALDELDDSLFVAAL	388
QY	358	ILISDGRGLVWRKTEIDIOJLLOALETOLKJLHNPESOLFAFLLOKMDLROIIVTEHO	417
Db	389	ICCGDRGLVWIPSTIEKMQESIYAVLKLHLOSNHPDPSLFPPALLOKMDLRLOIVTEHQA	448
QY	418	LLOYIKKTETDMSLHPLLOEIKDYK 443	
Db	449	LVOYIKKTETDMSLHPLLOEIKDYK 474	

```
Search completed: February 25, 2003, 03:59:00
Job time : 23 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2003, 03:56:01 ; Search time 38 seconds
(without alignments)
1740.539 Million cell updates/sec

Title: US-09-931-007A-1
Perfect score: 3508
Sequence: 1 MGFTLGDSPIDPESDSFDTL.....KTETDMLPLQLQETKDLX 688

Scoring table: BIOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2231	63.6	505	JC4859	peroxisome prolif
2	2160.5	61.6	504	JF0280	peroxisome prolif
3	2153	61.4	505	A54101	peroxisome prolif
4	2132	60.8	505	JC5777	peroxisome prolif
5	2043	58.2	475	JC4264	peroxisome prolif
6	2041	58.2	475	JF0279	peroxisome prolif
7	1471	41.9	477	C42214	peroxisome prolif
8	1100.5	31.4	441	A45360	steroid hormone re
9	1091.5	31.1	468	A45288	peroxisome prolif
10	1083.5	30.9	468	JC2085	peroxisome prolif
11	1082.5	30.9	440	I55442	peroxisome prolif
12	1076.5	30.7	440	JC4530	peroxisome prolif
13	1076.5	30.7	468	I56603	peroxisome prolif
14	1074.5	30.6	468	A49289	peroxisome prolif
15	1060.5	30.2	474	A42214	peroxisome prolif
16	946.5	27.0	396	B42214	peroxisome prolif
17	659	18.8	576	PC4290	peroxisome prolif
18	408	11.6	576	A57048	peroxisome prolif
19	397	11.3	578	S52913	nuclear receptor R
20	392	11.2	1394	B34598	nuclear receptor R
21	391	11.1	579	A57057	ecdysonone-induced p
22	385.5	11.0	1237	A34596	orphan nuclear hor
23	384.5	11.0	508	A30226	ecdysonone-induced p
24	379	10.8	614	A30268	thyroid/steroid ho
25	375.5	10.7	483	S05979	thyroid hormone re
26	371.5	10.6	1443	I505979	steroid hormone re
27	354.5	10.1	711	SA3464	ecdysonone-induced p
28	351	10.0	685	C36591	ecdysonone-induced p
29	341.5	9.7	468	A56856	E75 B steroid rece
					retinoid-related o

30	341.5	9.7	523	2	A53196	orphan hormone nuc
31	341.5	9.7	548	2	C53196	orphan hormone nuc
32	341.5	9.7	556	2	B53196	orphan hormone nuc
33	338.5	9.6	523	2	S68517	nuclear hormone re
34	331	9.4	448	2	B56558	retinoic acid rece
35	330	9.4	464	2	A56558	retinoic acid rece
36	328	9.4	459	2	A41977	retinoic acid rece
37	327.5	9.3	452	2	S78481	retinoic acid rece
38	327.5	9.3	458	2	S06123	retinoic acid rece
39	326.5	9.3	448	2	A43786	retinoic acid rece
40	326.5	9.3	445	2	S13512	retinoic acid rece
41	325.5	9.3	560	2	JC2494	nuclear receptor R
42	324.5	9.3	448	2	S02827	retinoic acid rece
43	323.5	9.2	444	2	I51256	retinoic acid rece
44	323.5	9.2	445	2	A56043	steroid hormone re
45	321.5	9.2	955	4	C40045	probable transcript

ALIGNMENTS

RESULT 1
JC4859
peroxisome proliferator-activated receptor gamma-2 - human
C:Species: Homo sapiens (man)
C:Date: 15-Aug-1996 #sequence_revision 18-Oct-1996 #text_change 20-Sep-1999
C:Accession: JC4859; S71351
R:Elbrecht, A.; Chen, Y.; Cullinan, C.A.; Hayes, N.; Lebowitz, M.D.; Moller, D.E.; B
Biochem. Biophys. Res. Commun. 224, 431-437, 1996
A>Title: Molecular cloning, expression and characterization of human peroxisome proli
A:Reference number: JC4859; MUID:96295505; PMID:8702406
A:Contents: fat
A:Accession: JC4859
A:Molecule type: mRNA
A:Residues: 1-505 <ELB>
A:Cross-references: GB:063415; NID:q1432176; PIDN:AAB04028.1; PID:q1432177
R:Rambe, K.G.; Tugwood, J.D.
Eur. J. Biochem. 239, 1-7, 1996
A>Title: A human peroxisome-proliferator-activated receptor-gamma is activated by ind
A:Reference number: S71351; MUID:96305559; PMID:8706692
A:Accession: S71351
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 29-505 <LAMB>
A:Cross-references: EMBL:X90563; NID:q1480099; PIDN:CA62152.1; PID:e258302; PID:g149
A:Experimental source: liver
C:Comment: This receptor plays a role in adipocyte differentiation and in regulating
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
C:Keywords: DNA binding; heterodimer; receptor; zinc finger
F:1-505/Product: peroxisome proliferator-activated receptor gamma-2 #status predicted
F:29-505/Product: peroxisome proliferator-activated receptor gamma-1 #status predicted
F:137-422/Domain: erba transforming protein homology <ERBA>

Query Match 63.6%; Score 2231; DB 2; Length 505;
Best Local Similarity 87.7%; Pred. No. 9.3e-117;
Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY	1	MGFTLGDSPIDPESDSFDTLSANISQENTWDTMPMPPTNFGISSVDLSWMEHSHSF	60
DB	1	MGFTLGDSPIDPESDSFDTLSANISQENTWDTMPMPPTNFGISSVDLSWMEHSHSF	60
QY	61	DKPPTVDFSSISPHYEDIPFTRTDPVADYKXDKLQEQSAIKVEPASPYYSEKT	120
DB	61	DKPPTVDFSSISPHYEDIPFTRTDPVADYKXDKLQEQSAIKVEPASPYYSEKT	120
QY	121	QLYN-----RKKCYCRQKCLAVGSMNAIRFGMPQAEKEKLLAEISSDIDQINPSSADLR	178
DB	121	QLYN-----RKKCYCRQKCLAVGSMNAIRFGMPQAEKEKLLAEISSDIDQINPSSADLR	178
QY	125	-----RKKCYCRQKCLAVGSMNAIRFGMPQAEKEKLLAEISSDIDQINPSSADLR	240
DB	181	RIHKSRNKKCYCRQKCLAVGSMNAIRFGMPQAEKEKLLAEISSDIDQINPSSADLR	240

QY 179 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMGEDIKIKFHHITPLOE 238
 |||||||
 Db 241 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMGEDIKIKFHHITPLOE 300
 QY 239 OSKEVAIRIFGCGFRSVAVOETFEYAKSIRGFVNLNDQYTLTKYGVHEIITMMLAS 298
 |||||||
 Db 301 OSKEVAIRIFGCGFRSVAVOETFEYAKSIRGFVNLNDQYTLTKYGVHEIITMMLAS 360
 QY 299 LMKKDVLLISEGCGFMTREFLKLRRKPGDFEMPRKFEFAVKFNALDSDLAIFAVII 358
 |||||||
 Db 361 LMKKDVLLISEGCGFMTREFLKLRRKPGDFEMPRKFEFAVKFNALDSDLAIFAVII 420
 QY 359 LSGDRPGLLVKPRFIEDIONLLOALELQKLNPRESSQLFAKLQKMTDLROIIVTEHVOL 418
 |||||||
 Db 421 LSGDRPGLLVKPRFIEDIONLLOALELQKLNPRESSQLFAKLQKMTDLROIIVTEHVOL 480
 QY 419 LOYIKKETDMSLHPLLOETIKDLY 443
 |||||||
 Db 481 LOYIKKETDMSLHPLLOETIKDLY 505

RESULT 2

JE0280
 peroxisome proliferator-activated receptor gamma2 - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C>Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 21-Jan-2000
 C:Accession: JE0280
 R:Griffiths, E.; Sundvold, H.; Klungland, H.; Lien, S.
 Biochem. Biophys. Res. Commun. 249, 713-718, 1998
 A:Title: Characterization of porcine peroxisome proliferator-activated receptors gamma1
 A:Reference number: JE0279; MUID:98401156; PMID:9731203
 A:Accession: JE0280
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-504 <GRD>
 A:Cross-references: GB:AJ006756
 C:Superfamily: unassigned erda-related proteins; erda transforming protein homology
 F:136-422/Domain: erda transforming protein homology <ERB>

Query Match 61.6%; Score 2160.5; DB 2; Length 504;
 Best Local Similarity 85.1%; Pred. No. 7.7e-113;
 Matches 430; Conservative 7; Mismatches 5; Indels 63; Gaps 2;
 QY 1 MGETLDSPIDESDSFDTLSANISQEMTWDTMPMPNFGISSVDLSVMDHSHSF 60
 |||||||
 Db 1 MGETLDSPIDESDSFDTLSANISQEMTWDTMPMPNFGISSVDLSVMDHSHSF 59
 QY 61 DIKPFVTVDSSISPHYEDIPRTDPVADYKYDKLOEQSAIKVPPASPYSEKT 120
 |||||||
 Db 60 DIKPFVTVDSSISPHYEDIPRTDPVADYKYDKLOEQSAIKVPPASPYSEKT 119
 QY 121 QLYN----- 124
 |||||
 Db 120 QLYNKRHEEPSNLMAIEKRVGDKASGFHYGVHACEGKGFRRIRIKLIYDRCDLNC 179
 QY 125 -----RNCQYCRROCKLAVGSHNAIRFGMPQAEKKLLAETSSDIDQNPESADLR 178
 |||||||
 Db 180 RIHKRSRNCQYCRROCKLAVGSHNAIRFGMPQAEKKLLAETSSDIDQNPESADLR 239
 QY 179 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMGEDIKIKFHHITPLOE 238
 |||||||
 Db 240 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMGEDIKIKFHHITPLOE 299
 QY 239 OSKEVAIRIFGCGFRSVAVOETFEYAKSIRGFVNLNDQYTLTKYGVHEIITMMLAS 298
 |||||||
 Db 300 OSKEVAIRIFGCGFRSVAVOETFEYAKSIRGFVNLNDQYTLTKYGVHEIITMMLAS 359
 QY 299 LMKKDVLLISEGCGFMTREFLKLRRKPGDFEMPRKFEFAVKFNALDSDLAIFAVII 358
 |||||||
 Db 360 LMKKDVLLISEGCGFMTREFLKLRRKPGDFEMPRKFEFAVKFNALDSDLAIFAVII 419
 QY 359 LSGDRPGLLVKPRFIEDIONLLOALELQKLNPRESSQLFAKLQKMTDLROIIVTEHVOL 418
 |||||||

Db 420 LSGDRPGLLVKPRFIEDIONLLOALELQKLNPRESSQLFAKLQKMTDLROIIVTEHVOL 479
 QY 419 LOYIKKETDMSLHPLLOETIKDLY 443
 |||||||
 Db 480 LOYIKKETDMSLHPLLOETIKDLY 504

RESULT 3

AS4101
 peroxisome proliferator-activated receptor gamma - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 20-Sep-1999
 C:Accession: AS4101; A57740; A49294; JN0881; S53748
 R:Tontoz, P.; Hu, E.; Graves, R.A.; Budavari, A.I.; Spiegelman, B.M.
 Genes Dev. 8, 1224-1234, 1994
 A:Title: mPARGamma2: tissue-specific regulator of an adipocyte enhancer.
 A:Reference number: AS4101; MUID:95011536; PMID:7926726
 A:Accession: AS4101
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-505 <TON>
 A:Cross-references: EMBL:U09138; NID:9500639; PIDN:AAA62277.1; PID:9500640
 A:Experimental source: adipose tissue
 R:Kliwer, S.A.; Forman, B.M.; Blumberg, B.; Ong, E.S.; Borgmeyer, U.; Mangelsdorf, D.
 Proc. Natl. Acad. Sci. U.S.A. 91, 7355-7359, 1994
 A:Title: Differential expression and activation of a family of murine peroxisome prol
 A:Reference number: A57740; MUID:94316694; PMID:8041794
 A:Accession: A57740
 A:Molecule type: mRNA
 A:Residues: 31-382, 'S', 384-505 <KLI>
 A:Cross-references: GB:U10374; NID:9507776; PIDN:AAA19971.1; PID:9514307; GB:U10375
 A:Experimental source: liver
 R:Zhu, Y.; Alvares, K.; Huang, Q.; Rao, M.S.; Reddy, J.K.
 J. Biol. Chem. 268, 26817-26820, 1993
 A:Title: Cloning of a new member of the peroxisome proliferator-activated receptor ge
 A:Reference number: A49294; MUID:94086482; PMID:8262913
 A:Accession: A49294
 A:Molecule type: mRNA
 A:Residues: 31-505 <ZHU>
 A:Cross-references: GB:U01841; NID:9454000; PIDN:AA622134.1; PID:9454001
 A:Experimental source: liver
 R:Chen, F.; Law, S.W.; O'Malley, B.W.
 Biochem. Biophys. Res. Commun. 196, 671-677, 1993
 A:Title: Identification of two mPAP related receptors and evidence for the existence
 A:Reference number: JN0881; MUID:94059089; PMID:8240342
 A:Accession: JN0881
 A:Molecule type: mRNA
 A:Residues: 31-212, 'DR', 215-280, 'S', 282, 'F', 284-382, 'S', 384-441, 447-496, 'F', 498-505 <
 A:Experimental source: heart
 R:Tontoz, P.; Graves, R.A.; Budavari, A.I.; Erdjument-Bromage, H.; Lui, M.; Hu, E.;
 Nucleic Acids Res. 22, 5628-5634, 1994
 A:Title: Adipocyte-specific transcription factor ARF6 is a heterodimeric complex of t
 A:Reference number: S53748; MUID:95140625; PMID:7838715
 A:Accession: S53748
 A:Molecule type: Protein
 A:Residues: 66-85; 146-160 <TOW>
 A:Experimental source: adipose tissue
 C:Complex: heterodimer; peroxisome proliferator-activated receptor gamma and retinoid
 C:Function:
 A:Description: transcription regulation of adipogenesis; activates transcription in r
 A>Note: binding of the heterodimer to its recognition element induces DNA bending
 C:Superfamily: unassigned erda-related proteins; erda transforming protein homology
 C:Keywords: DNA binding; heterodimer; nucleus; receptor; transcription regulation; z1
 F:137-423/Domain: erda transforming protein homology <ERBA>
 F:139-159/Region: zinc finger CCCC motif
 F:160-224/Region: DNA binding #status predicted
 F:176-198/Region: zinc finger CCCC motif

Query Match 61.4%; Score 2153; DB 2; Length 505;
 Best Local Similarity 84.0%; Pred. No. 2e-112;
 Matches 424; Conservative 9; Mismatches 10; Indels 62; Gaps 1;
 QY 1 MGETLDSPIDESDSFDTLSANISQEMTWDTMPMPNFGISSVDLSVMDHSHSF 60

```

Db      1  MGETLGDSVPDDEHCAFDALPMSTSQEITMVDTEMPTWPNFGISSVDLSVMEDHSF 60
Oy      61  DIKPTTYDFFSSISPHNEDIPTFTDVPVADYKDKDLQEOYSAIKPEPASPYSSEKT 120
Db      61  DIKPTTYDFFSSISAPHNEDIPTFTADPMVADYKDKDLQEOYSAIKPEPASPYSSEKT 120
Oy      121  QLYN----- 124
Db      121  QLYNRPHEEPSNLMAIECRVGDKASGFHYGVNAHECGCKGFRTTIRLKLHYRCDLNC 180
Oy      125  -----RKKQYCRQKQKLAAGMSHNAIRFGMRPAEKEKLAEISSDIDLNPSADLR 178
Db      181  RIHKRSNRKCOYCRQKQKLAAGMSHNAIRFGMRPAEKEKLAEISSDIDLNPSADLR 240
Oy      179  ALAKLHYDSYIKSPFLTKAKARAILTGKTTDPSPIVYIDMSLMNGEIKFKHTHPLOE 238
Db      241  ALAKLHYDSYIKSPFLTKAKARAILTGKTTDPSPIVYIDMSLMNGEIKFKHTHPLOE 300
Oy      239  OSKEVALRIFQCGQFRSVEAVQEIETEVAKSIPGFVNLDLNDQVTLTKGYVHEIITYMLAS 298
Db      301  OSKEVALRIFQCGQFRSVEAVQEIETEVAKNIPGFINLNDQVTLTKGYVHEIITYMLAS 360
Oy      299  LMNKGVLISQGGWMTDEFLKSLRKPFGDMPEKFEFAVAFNALDSDLAIFIAVII 358
Db      361  LMNKGVLISQGGWMTDEFLKSLRKPFGDMPEKFEFAVAFNALDSDLAIFIAVII 420
Oy      359  LSGDRPGLLNYPIDIDIDNLLQALELOKLNHPPSSLOFLKLOKMTDLROIVTEHQL 418
Db      421  LSGDRPGLLNYPIDIDIDNLLQALELOKLNHPPSSLOFLKLOKMTDLROIVTEHQL 480
Oy      419  LQVTKTETDMSLHPLQEIYKDLV 443
Db      481  LHVIKTETDMSLHPLQEIYKDLV 505

RESULT 4
JC5777
peroxisome proliferator-activated receptor gamma 2 - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 24-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 24-Sep-1999
C:Accession: JC5777
A:Biochem. Biophys. Res. Commun. 239, 857-861, 1997
A:Title: Characterization of bovine peroxisome proliferator-activated receptors gamma 1
A:Reference number: JC5777; MUID:98042483; PMID:9367859
A:Accession: JC5777
A:Molecule type: mRNA
A:Residues: 1-505 <SD>
A:Cross-references: GB:Y12420; NID:92653405; PIDN:CAA7303.1; PID:ai188692; PID:92653406
A:Note: The sequences of residues 301-320 and 321-340 are interchanged in the authors'
hondrial beta-oxidation, and adipocyte differentiation.
C:Genetics:
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
F:1-505/Product: peroxisome proliferator-activated receptor gamma 2 #status predicted
F:31-505/Product: peroxisome proliferator-activated receptor gamma 1 #status predicted
F:137-423/Domain: erba transforming protein homology <ERBA>

Query Match 60.8%; Score 2132; DB 2; Length 505;
Best Local Similarity 83.6%; Pred. No. 2,9e-111;
Matches 422; Conservative 11; Mismatches 10; Indels 62; Gaps 1;

Oy      1  MGETLGDSVPDDEHCAFDALPMSTSQEITMVDTEMPTWPNFGISSVDLSVMEDHSF 60
Db      1  MGETLGALIDPESEPAFTVSARTSOEITWDEMPWPNFGISSVDLSMDHSHAF 60
Oy      61  DIKPTTYDFFSSISPHNEDIPTFTDVPVADYKDKDLQEOYSAIKPEPASPYSSEKT 120
Db      61  DIKPTTYDFFSSISPHNEDIPTFTADPMVADYKDKDLQEOYSAIKPEPASPYSSEKT 120
Oy      121  QLYN----- 123

```

Db	121	QLYSKPHEEPPNSLMAIEECVCCGKASGFHYGVACEGCKGFRRTIRLKLTYRCOLNC	180
QY	124	-----NRKKCCYCFQKCLAVGSMHNAIRFGMPAEREKELIAESSDIDOLNESPADLR	178
Db	181	RHHKSRNKKCOYCRQKCLAVGSMHNAIRFGMPAEREKELIAESSDIDOLNESPADLR	240
QY	179	ALAKHLNDYSYTKSPFLTKAKARAILTGKTTDKSPFVIYDMSLWMDGEDIKFKHTPLQOE	238
Db	241	ALAKHLNDYSYTKSPFLTKAKARAILTGKTTDKSPFVIYDMSLWMDGEDIKFKHTPLQOE	300
QY	239	OSKEVAIRIFGCGCFRSEVAEOETTEYAKSIJBGVNLNDQVTLKGVHEIITYTLAS	298
Db	301	PSKEVAIRIFGCGCFRSEVAEOETTEYAKNIPGFNLNDQVTLKGVHEIITYTLAS	360
QY	299	LMNKGVLISSEGGFMTEPEFLSKLRKPGDMEPEYFEFEAVKFNMLEDDSLAIFIAVII	358
Db	361	LMNKGVLISSEGGFMTEPEFLSKLRKPGDMEPEYFEFEAVKFNMLEDDSLAIFIAVII	420
QY	359	LSGDRPGLLNKPRIDIDNLQALELOLKLNHPESSOLFAKLLQKMTDLQOYTEHVOL	418
Db	421	LSGDRPGLLNKPRIDIDNLQALELOLKLNHPESSOLFAKLLQKMTDLQOYTEHVOL	480
QY	419	LOVYKKTETDMSLHPLDQIYKDLV 443	
Db	481	LOVYKKTETDMSLHPLDQIYKDLV 505	
RESULT 5			
	JC4264	peroxisome proliferator-activated receptor gamma - Chinese hamster	
	C:Species:	Cricetulus griseus (Chinese hamster)	
	C:Date:	15-oct-1995 #sequence_revision 08-Feb-1996 #text_change 01-Dec-2000	
	C:Accession:	JC4264; S42849	
	R:Aperlo, C.; Pognonec, P.; Saladin, R.; Auwerx, J.; Boulikos, K.E.		
	Gene 162, 297-302, 1995		
	A:Title:	cDNA cloning and characterization of the transcriptional activities of the ha	
	A:Reference number:	JC4264; MUID:96032361; PMID:7557447	
	A:Accession:	JC4264	
	A:Molecule type:	mRNA	
	A:Residues:	1-475 <AB>	
	A:Cross-references:	EMBL:Z30972; NID:9461358; PIDN:CAA83219.1; PID:9461359	
	C:Superfamily:	unassigned erba-related proteins; erba transforming protein homology	
	C:Keywords:	DNA binding; peroxisome; steroid hormone receptor; transcription regulati	
	F:107-399/Domain:	erba transforming protein homology <ERBA>	
Query Match 58.2%; Score 2043; DB 2; Length 475;			
Best Local Similarity 84.8%; Pred. No. 2.4e-106;			
Matches 403; Conservative 6; Mismatches 4; Indels 62; Gaps 1;			
QY	31	MVDTEMPEMPNPFNGISSVDLSVMEDHSHSPDKPPTTYDFSSISPHYEDIPFTRTDPVY	90
Db	1	MVDTEMPEMPNPFNGISSVDLSVMDDHSHSPDKPPTTYDFSSISAPHYEDIPFTRADPMV	60
QY	91	ADYKYDLKLOEQYSAIKYEPASPPYSEKTLQYN-----	124
Db	61	ADYKYDLKLOEQYSAIKYEPASPPYSEKADQLYNNPHEEPSNSLMAIECRVCGKASGFH	120
QY	125	-----NRKKCOYCFQKCLAVGSMHNAIRF	148
Db	121	YGVHACEGCKGFRRTIRLKLTYRCOLNCRILHKKSRNKKCOYCRQKCLAVGSMHNAIRF	180
QY	149	GMPAEREKELIAESSDIDOLNESPADLRALKHLYSYTKSPFLTKAKARAILTGKTT	208
Db	181	GMPAEREKELIAESSDIDOLNESPADLRALKHLYSYTKSPFLTKAKARAILTGKTT	240
QY	209	DKSPVIYDMSLWMDGEDIKFKHTPLQOESKEVAIRIFGCGCFRSEVAEOETTEYAKS	268
Db	241	DKSPVIYDMSLWMDGEDIKFKHTPLQOESKEVAIRIFGCGCFRSEVAEOETTEYAKN	300
QY	269	IPGFNLNDQVTLKGVHEIITYTLASLMNKGVLISSEGGFMTEPEFLSKLRKPGD	328
Db	301	IPGFNLNDQVTLKGVHEIITYTLASLMNKGVLISSEGGFMTEPEFLSKLRKPGD	360

QY 329 FMEPKFEFAVKNALDELDDSDLAIFAVIITLSGDRPGLNVKPIEDIQDNLQALELQK 388
 |||||||
 Db 361 FMEPKFEFAVKNALDELDDSDLAIFAVIITLSGDRPGLNVKPIEDIQDNLQALELQK 420
 |||||||
 QY 389 LNHPESSQLFAKLQKMDLRQIVTEHVOLLVYIKKFTDMSLHPLLOEYKDL 443
 |||||||
 Db 421 LNHPESSQLFAKLQKMDLRQIVTEHVOLLVYIKKFTDMSLHPLLOEYKDL 475
 |||||||

RESULT 6
 JF0279
 Peroxisome proliferator-activated receptor gamma 1 - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 21-Jan-2000
 C:Accession: JF0279
 R:Grindflek, E.; Sundvold, H.; Klungland, H.; Lien, S.
 Biochem. Biophys. Res. Commun. 249, 713-718, 1998
 A>Title: Characterization of porcine peroxisome proliferator-activated receptors gamma1
 A:Reference number: JF0279; MUID:98401156; PMID:9731203
 A:Accession: JF0279
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-475 <GR>
 A:Cross-references: GB:A006758
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 F:107-393/Domain: erba transforming protein homology <ERB>

Query Match 58.2%; Score 2041; DB 2: Length 475;
 Best Local Similarity 85.1%; Pred. No. 3,1e-106;
 Matches 404; Conservative 5; Mismatches 4; Indels 62; Gaps 1;

QY 31 MVDTEMPFWPT-NFGISSVDLSVMDHSHFDIKPFTTVDFSSISTPHYEDIPFTTRDPVY 90
 |||||||
 Db 1 MVDTEMPFWPTNFGISSVDLSVMDHSHFDIKPFTTVDFSSISTPHYEDIPFTTRDPVY 60
 |||||||

QY 91 ADYKDYDLQEOYSAIKVEPASPYYSEKTOLYN----- 124
 |||||||
 Db 61 ADYKDYDLQEOYSAIKVEPASPYYSEKTOLYNKRHEPSSNLMAIECRVCGDKASGFH 120
 |||||||

QY 125 -----RNKCOYCRFOKCLAVGMSHNAIRF 148
 |||||||

Db 121 YGVNAHCEGCKGFRRRTIRKLILYDRCDLNCRIHKSSRNKCQYCRFOKCLAVGMSHNAIRF 180
 |||||||

QY 149 GRMPOAEKELKLAETSSDIDQNPESADRLAKHLVDSYIKSPFLTKAKARAILTGKTT 208
 |||||||
 Db 181 GRMPOAEKELKLAETSSDIDQNPESADRLAKHLVDSYIKSPFLTKAKARAILTGKTT 240
 |||||||

QY 209 DKSPFVIYDMSLMNGEDKIKFKHITPLQEOSEKVAIRIFGCGFRSVAVOETETAKS 268
 |||||||
 Db 241 DKSPFVIYDMSLMNGEDKIKFKHITPLQEOSEKVAIRIFGCGFRSVAVOETETAKN 300
 |||||||

QY 269 IPGFVNLNDLNDQVTLTKYVHEIITYMLASLNNKGVLSSEGQFWTRFLSKLRKPRD 328
 |||||||
 Db 301 IPGFVNLNDLNDQVTLTKYVHEIITYMLASLNNKGVLSSEGQFWTRFLSKLRKPRD 360
 |||||||

QY 329 FMEPKFEFAVKNALDELDDSDLAIFAVIITLSGDRPGLNVKPIEDIQDNLQALELQK 388
 |||||||
 Db 361 FMEPKFEFAVKNALDELDDSDLAIFAVIITLSGDRPGLNVKPIEDIQDNLQALELQK 420
 |||||||

QY 389 LNHPESSQLFAKLQKMDLRQIVTEHVOLLVYIKKFTDMSLHPLLOEYKDL 443
 |||||||
 Db 421 LNHPESSQLFAKLQKMDLRQIVTEHVOLLVYIKKFTDMSLHPLLOEYKDL 475
 |||||||

RESULT 7
 C42214
 Peroxisome proliferator-activated receptor gamma chain - African clawed frog
 N:Alternate names: PPAR-gamma
 C:Species: Xenopus laevis (African clawed frog)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
 C:Accession: C42214
 R:Dreyer, C.; Krey, G.; Keller, H.; Givel, F.; Helftenbein, G.; Wahli, W.

Cell 68, 879-887, 1992
 A>Title: Control of the peroxisomal beta-oxidation pathway by a novel family of nucle
 A:Reference number: A42214; MUID:92191267; PMID:1312391
 A:Accession: C42214
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-477 <DR>
 A:Cross-references: GB:M84163; NID:9214667; PID:AAA49937.1; PID:9214668
 A>Note: sequence extracted from NCBI backbone (NCBI:88295)
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: DNA binding; transcription regulation; zinc finger
 F:111-395/Domain: erba transforming protein homology <ERBA>

Query Match 41.9%; Score 1471; DB 2: Length 477;
 Best Local Similarity 62.7%; Pred. No. 1.4e-74;
 Matches 307; Conservative 44; Mismatches 49; Indels 90; Gaps 9;

QY 31 MVDTEMPFWPT-NFGISSVDLSVMDHSHFDIKPFTTVDFSSISTPHYEDIP-----FT 84
 |||||||
 Db 1 MVDTEMPFWPTNFGISSVDLSVMDHSHFDIKPFTTVDFSSISTPHYEDIP-----FT 59
 |||||||

QY 85 RNDPVVADYKDYDLQEOYSAIKVEPASPYYSEKTO----- 121
 |||||||

Db 60 RNDQSPIDYKDYDLQEOYSAIKVEPASPYYSEKTO----- 119
 |||||||

QY 122 -----LY-----NNKCOYCRFOKCLAVGMSH 143
 |||||||

Db 120 ASGFHYGVNAHCEGCKGFRRRTIRKLILYDRCDLNCRIHKSSRNKCQYCRFOKCLAVGMSH 179
 |||||||

QY 144 NAIRFGRMPOAEKELKLAETSSDIDQNPESADRLAKHLVDSYIKSPFLTKAKARAIL 203
 |||||||

Db 180 NAIRFGRMPOAEKELKLAETSSDIDQNPESADRLAKHLVDSYIKSPFLTKAKARAIL 239
 |||||||

QY 204 TPKT---TKSPFVIYDMSLMNGEDKIKFKHITPLQEO-----SKEVAIRIFGCGF 253
 ||:::|

Db 240 DQSHRQNSRGYTRHGLDGDGSDQGAVER--PRAEGGCGDSNLPALVALR----- 290
 ||:::|

QY 254 RSEVAOETETAKSIPGFVNLNDLNDQVTLTKYVHEIITYMLASLNNKGVLSSEGQF 313
 ||:::|

Db 291 ---GGVRETFERAKNIPGVSLDNDQVTLTKYVHEIITYMLASLNNKGVLSSEGQF 347
 ||:::|

QY 314 MTRFELSLRKPFDFMEPKFEFAVKNALDELDDSDLAIFAVIITLSGDRPGLNVKPIE 373
 |||||||

Db 348 MTRFELSLRKPFDFMEPKFEFAVKNALDELDDSDLAIFAVIITLSGDRPGLNVKPIE 407
 |||||||

QY 374 DIQDNLQALELQKMDLRQIVTEHVOLLVYIKKFTDMSLHPLLOEYKDL 433
 |||||||

Db 408 DIQDNLQALELQKMDLRQIVTEHVOLLVYIKKFTDMSLHPLLOEYKDL 467
 |||||||

QY 434 LLOEYKDL 443
 |||||||

Db 468 LLOEYKDL 477
 |||||||

RESULT 8
 A45360
 Steroid hormone receptor delta - human
 N:Alternate names: nuclear hormone receptor 1 (NUC1); peroxisome proliferator-activat
 C:Species: Homo sapiens (man)
 C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 31-Oct-1997
 C:Accession: A45360
 R:Schmidt, A.; Endo, N.; Rutledge, S.J.; Vogel, R.; Shinar, D.; Rodan, G.A.
 Mol. Endocrinol. 6, 1634-1641, 1992
 A>Title: Identification of a new member of the steroid hormone receptor superfamily t
 A:Reference number: A45360; MUID:93078797; PMID:1333051
 A:Accession: A45360
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: nucleic acid
 A:Residues: 1-441 <SCH>
 A:Experimental source: osteosarcoma SAOS-2/B10 cells
 A>Note: sequence extracted from NCBI backbone (NCBI:118801)
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
 C:Keywords: DNA binding; steroid hormone receptor; transcription regulation; zinc fin


```

QY 63 KPF-----TIVDFSSISTPHXEDIPETFTDPVADYKYDLKLOEYQSAIKVE 109
Db 52 ADYQYLGSCPGSEGSVITDTLSPASSPSSVSCPVI--PASTD-----ESPGSALNIE 101
QY 110 -----PASPPV-----SEKTOLYNNRKCQYCPHQC 136
Db 102 CRICGDKAGYHGVACAGCGCKFFRRTRRLKLYDKCRSCRTQCKNNRKCQYCFPHQC 161
QY 137 LAVGSMHNAIRFGMPQAEKEKELAEI-SSDIDQLNPESADLRALAKHLVDSYIKSFPLT 195
Db 162 LSVGSMHNAIRFGMRPSRKAKLKAETILCEHDLKDSERADLSLGRKHAEVATKNFNNN 221
QY 196 KAKARAILTGKTTDKSPPEYIYDMSLMDGEDKIKFKHITPLOSQSEVAIRIFQSCQFRS 255
Db 222 KYKARVILLAGKTSNNRPFYIHMETLCMAEKTLYAMKANGAVE-DKEAEVREFHCCQCS 280
QY 256 VEAVOETTEVAKSIPGFVNLNDQVTLTKYGHETIYMLASIMKDDVLLSEGGFMT 315
Db 281 VETVTELFEPKAIPEGFANLDLDQVTLTKYGEAIFPMLSLMMKDKGLIAYGNGEFT 340
QY 316 REFLLSRKPFQDMEKFEFAVFNALDELDDSLAIFTAVIILSGDRPGLLVKPIEDI 375
Db 341 REFLLKLRKPFQDIMEKFKDFAMKFNALDELDDSLIFVAIITCGDRPGLNIGYIEKL 400
QY 376 QDNLLQALELQALKNHPESQLEFAKLQKMTDLQIYVTEHVOLLQYIKKTFDMSLHPL 435
Db 401 QEGIVVTLKHLQSNHDPDTFLFKLLQKRVLDLQIYVTEHQAOLVQYIKKTESDALHPLL 460
QY 436 QETIKDLY 443
Db 461 QETIKDLY 468

RESULT 11
155442
peroxisome proliferator-activated receptor delta chain - mouse
N:Alternate names: peroxisome proliferator-activated receptor NUC1
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Sep-1999
C:Accession: I55442; B57740; PNO676
R:Amli, E.Z.; Bonino, F.; Allnaud, G.; Abumrad, N.A.; Grimaldi, P.A.
J. Biol. Chem. 270, 2367-2371, 1995
A:Title: Cloning of a protein that mediates transcriptional effects of fatty acids in pr
A:Reference number: I55442; MUID:95138211; PMID:7836471
A:Accession: I55442
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-440 <RES>
A:Cross-references: GB:L28116; NID:9452113; PIDN:AAA63394.1; PID:9452114
R:Kliwer, S.A.; Forman, B.M.; Blumberg, B.; Ong, E.S.; Borgmeyer, U.; Mangelsdorf, D.J.
Proc. Natl. Acad. Sci. U.S.A. 91, 7355-7359, 1994
A:Title: Differential expression and activation of a family of murine peroxisome prolif
A:Reference number: A57740; MUID:94316694; PMID:8041794
A:Accession: B57740
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-148, 'DG', 151-440 <KL1>
A:Cross-references: GB:U10374; GB:U10375; NID:9507778; PIDN:AAA19972.1; PID:9514308
R:Chen, F.; Law, S.W.; O'Malley, B.W.
Biochem. Biophys. Res. Commun. 196, 671-677, 1993
A:Title: Identification of two mPPAR related receptors and evidence for the existence of
A:Reference number: JN0881; MUID:94059089; PMID:8240342
A:Accession: PNO676
A:Molecule type: mRNA
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 73-137 <CHE>
A:Experimental source: brain
C:Comment: This protein has function of peroxisome proliferator activated receptor.
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
C:Keywords: DNA binding; hormone receptor; nucleus; zinc finger
F:71-358/Domain: erba transforming protein homology <ERBA>
F:73-136/Region: DNA binding #status predicted

```

```

Query Match 30.9%; Score 1082.5; DB 2; Length 440;
Best Local Similarity 62.6%; Pred. No. 4.9e-53;
Matches 204; Conservative 59; Mismatches 62; Indels 1; Gaps 1;

QY 119 KQOLYNNRKCQYCRPOKCLAVGSMHNAIRFGMPQAEKEKELAEI-SSDIDQLNPESADL 177
Db 115 KIQKNNRKCQYCRPOKCLAVGSMHNAIRFGMPQAEKEKELAEI-SSDIDQLNPESADL 174
QY 178 RALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPPEYIYDMSLMDGEDKIKFKHITPQ 237
Db 175 KAFSKHIIYAIYALKNFMTKKARSLITGKSSHNAIPFVIDIETLQAEKGLWVKQLVN 234
QY 238 EOSKEVAIRIFQSCQFRSVEAVOETTEVAKSIPGFVNLNDQVTLTKYGHETIYMLA 297
Db 235 PYNNEISVHVFRCSTVETVRELFEPKAIPEGFANLDLDQVTLTKYGEAIFAMLA 294
QY 298 SLMNNDGVILISGCGFMTEFLSLRKPRGDMPEKFEFAVFNALDELDDSLAIFTAVI 357
Db 295 SLVNNKDGILLVANGSGFVHEFLRSRKPPSDIIEPFERFAVFNALDELDDSLAIFTAVI 354
QY 358 IISGDRPGLLVKPIEDIDQNLQALELQALKNHPESQLEFAKLQKMTDLQIYVTEHVQ 417
Db 355 IICGDRPGLLVNPPQYEAIDDTILRALFEHLQVNHDPDTFLFKLLQKRVLDLQIYVTEHQA 414
QY 418 LLQVYIKKTFDMSLHPLLQETIKDLY 443
Db 415 MNQWLKTTSEFTLHPLLQETIKDLY 440

RESULT 12
JC4530
peroxisome proliferator activated receptor delta chain - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 16-Apr-1996 #sequence_revision 24-May-1996 #text_change 20-Sep-1999
C:Accession: JC4530
R:Xing, G.; Zhang, L.; Zhang, L.; Heynen, T.; Yoshikawa, T.; Smith, M.; Weiss, S.; De
Biochem. Biophys. Res. Commun. 217, 1015-1025, 1995
A:Title: Rat PPAR delta contains a CGG triplet repeat and its prominently expressed 1
A:Reference number: JC4530; MUID:96125342; PMID:8554552
A:Accession: JC4530
A:Molecule type: mRNA
A:Residues: 1-440 <XIN>
A:Cross-references: GB:U40064; NID:g1185466; PIDN:AAC5419.1; PID:g1185467
A:Experimental source: brain
C:Comment: This receptor modulates CNS function by regulating the transcriptional act
t is involved in the response to thermal and pain sensations or cold adaptation and n
C:Genetics:
A:Gene: ppar delta
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
C:Keywords: brain; DNA binding; peroxisome; receptor; zinc finger
F:71-358/Domain: erba transforming protein homology <ERBA>
F:73-938/Region: zinc finger CCCC motif
F:110-127/Region: zinc finger CCCC motif
F:245-440/Domain: ligand binding #status predicted <LIB>

Query Match 30.7%; Score 1076.5; DB 2; Length 440;
Best Local Similarity 62.0%; Pred. No. 1.1e-52;
Matches 202; Conservative 60; Mismatches 63; Indels 1; Gaps 1;

QY 119 KQOLYNNRKCQYCRPOKCLAVGSMHNAIRFGMPQAEKEKELAEI-SSDIDQLNPESADL 177
Db 115 KIQKNNRKCQYCRPOKCLAVGSMHNAIRFGMPQAEKEKELAEI-SSDIDQLNPESADL 174
QY 178 RALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPPEYIYDMSLMDGEDKIKFKHITPQ 237
Db 175 KAFSKHIIYAIYALKNFMTKKARSLITGKSSHNAIPFVIDIETLQAEKGLWVKQLVN 234
QY 238 EOSKEVAIRIFQSCQFRSVEAVOETTEVAKSIPGFVNLNDQVTLTKYGHETIYMLA 297
Db 235 PYNNEISVHVFRCSTVETVRELFEPKAIPEGFANLDLDQVTLTKYGEAIFAMLA 294
QY 298 SLMNNDGVILISGCGFMTEFLSLRKPRGDMPEKFEFAVFNALDELDDSLAIFTAVI 357
Db 295 SLVNNKDGILLVANGSGFVHEFLRSRKPPSDIIEPFERFAVFNALDELDDSLAIFTAVI 354

```

Db 295 STVNKDGILLVANGSGFVTHEFLRSIRKPSDIIIEPKFEFAVFNALVELVSDLALEIAI 354
QY 358 IISGDRPGLLNKPIEDIDIDNLLQALEQLKLNHPESQSLFPAKLQKMTDLROIVTEHQ 417
Db 355 IICGDRPGMLNVPQVEAIDDTLLQALEFHLQVNHDPQSYLFPKLLQKMDLRQLVTEHAQ 414
QY 418 LLQVIKRTETDMSLHPLLOEIYKDY 443
Db 415 MMQWLKKTESFLLHPLLOEIYKDY 440

RESULT 13
156603
peroxisome proliferator activated receptor alpha - human
C:Species: Homo sapiens (man)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Sep-1999
C:Accession: I56603
R:Murkhetjee, R.; Jow, L.; Noonan, D.; McDonnell, D.P.
J. Steroid Biochem. Mol. Biol. 51, 157-166, 1994
A:Title: Human and rat peroxisome proliferator activated receptors (PPARs) demonstrate s
A:Reference number: I56603; MUID:95071923; PMID:7981125
A:Accession: I56603
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-468 <RES>
A:Cross-references: GB:S74349; NID:9765239; PIDN:AA832649.1; PID:9765240
C:Genetics:
A:Gene: hPPARalpha
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
C:Keywords: zinc finger
F:100-386/Domain: erba transforming protein homology <ERBA>

Query Match 30.7%; Score 1076.5; DB 2; Length 468;
Best Local Similarity 49.8%; Pred. No. 1.2e-52;
Matches 236; Conservative 65; Mismatches 106; Indels 67; Gaps 9;

QY 31 MVDTEMPMPN-----FGISSVDSLVMEDSHSFDIKPFTTV--- 68
Db 1 MVDTEPLCLPLSLBAGDLESPLEEFLOEMNIOEISQISGEDSSGSGFTGYGLGSC 60
QY 69 ---DSSIS-TPHYEDIPTRTDPVADYKYLQOEYOS-AIKVE-----PASPPY- 116
Db 61 PGSDSVITDITLSPASSPSVTPVPG-----SVDESPGALNIECRICGDKASGYHNG 115
QY 117 -----SEKTYLNKKCOYCFQFCOCLAVGSMHNAIRGR 150
Db 116 VIACGCGKFFRTTRTLKLYDKCDRCKIQKNNKCOYCFHKLCSGSMHNAIRGR 175
QY 151 MPOAEKELIAEI-SSDIDQNPESADLRALAKHLYDSYKSPPLTKAKARAILTGKTTD 209
Db 176 MPRSEKAKIAELILCEHIEDSEFADLSLAKRIYEAVLKNFNKNKVKARVILSGKASN 235
QY 210 KSPFYIYDMSNLMGEDKTKFKHITPLOSKEVAIRIFGCGFVSVEAVQETITEYAKSI 269
Db 236 NPPFYIHMETLMAEKTLYAK-LVANGIONKEAEVRIFHCOCQSVETVTELTFAKAI 294
QY 270 PGFVNLDLNDQYTLKYGVEHIIYTMASLMMKDGVLSEGGFMTREFLKLRRPGDF 329
Db 295 PGFANLDLNDQYTLKYGVEHIFAMLSVMMKDGDLVAVGGFTRREFLKLRRPGDF 354
QY 330 MBPKREFAVFNALVELDSDLAIFLAVIILSGDRPGLNVKPIEDIDONLLOALELOKL 389
Db 355 MBPKREFAVFNALVELDSDISLFAVAIICGDRPGLNVGHIEKMGESIVAVLRLHQS 414
QY 390 NIPESQSLFPAKLQKMTDLROIVTEHVOLLQYIKTETDMSLHPLLOEIYKDY 443
Db 415 NHPDDIFLFPKLLQKMDLRQLVTEHAQVLQIKKTESDALHPLLOEIYRDMY 468

RESULT 14
A49289
peroxisome proliferator-activated receptor - human
C:Species: Homo sapiens (man)

C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 20-Sep-1999
C:Accession: A49289
R:Sher, T.; Yi, H.F.; McBride, O.W.; Gonzalez, F.J.
Biochemistry 32, 5598-5604, 1993
A:Title: cDNA cloning, chromosomal mapping, and functional characterization of the hu
A:Reference number: A49289; MUID:93277839; PMID:7684926
A:Accession: A49289
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-468 <SHD>
A:Cross-references: GB:L02932; NID:9307340; PIDN:AAA36468.1; PID:9307341
A>Note: authors translated the codon AAC for residue 33 as Asp
C:Genetics:
A:Map position: 22
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
C:Keywords: DNA binding; transcription regulation; zinc finger
F:100-386/Domain: erba transforming protein homology <ERBA>

Query Match 30.6%; Score 1074.5; DB 2; Length 468;
Best Local Similarity 49.8%; Pred. No. 1.5e-52;
Matches 236; Conservative 65; Mismatches 106; Indels 67; Gaps 9;

QY 31 MVDTEMPMPN-----FGISSVDSLVMEDSHSFDIKPFTTV--- 68
Db 1 MVDTEPLCLPLSLBAGDLESPLEEFLOEMNIOEISQISGEDSSGSGFTGYGLGSC 60
QY 69 ---DSSIS-TPHYEDIPTRTDPVADYKYLQOEYOS-AIKVE-----PASPPY- 116
Db 61 PGSDSVITDITLSPASSPSVTPVPG-----SVDESPGALNIECRICGDKASGYHNG 115
QY 117 -----SEKTYLNKKCOYCFQFCOCLAVGSMHNAIRGR 150
Db 116 VIACGCGKFFRTTRTLKLYDKCDRCKIQKNNKCOYCFHKLCSGSMHNAIRGR 175
QY 151 MPOAEKELIAEI-SSDIDQNPESADLRALAKHLYDSYKSPPLTKAKARAILTGKTTD 209
Db 176 MPRSEKAKIAELILCEHIEDSEFADLSLAKRIYEAVLKNFNKNKVKARVILSGKASN 235
QY 210 KSPFYIYDMSNLMGEDKTKFKHITPLOSKEVAIRIFGCGFVSVEAVQETITEYAKSI 269
Db 236 NPPFYIHMETLMAEKTLYAK-LVANGIONKEAEVRIFHCOCQSVETVTELTFAKAI 294
QY 270 PGFVNLDLNDQYTLKYGVEHIIYTMASLMMKDGVLSEGGFMTREFLKLRRPGDF 329
Db 295 PGFANLDLNDQYTLKYGVEHIFAMLSVMMKDGDLVAVGGFTRREFLKLRRPGDF 354
QY 330 MBPKREFAVFNALVELDSDLAIFLAVIILSGDRPGLNVKPIEDIDONLLOALELOKL 389
Db 355 MBPKREFAVFNALVELDSDISLFAVAIICGDRPGLNVGHIEKMGESIVAVLRLHQS 414
QY 390 NIPESQSLFPAKLQKMTDLROIVTEHVOLLQYIKTETDMSLHPLLOEIYKDY 443
Db 415 NHPDDIFLFPKLLQKMDLRQLVTEHAQVLQIKKTESDALHPLLOEIYRDMY 468

RESULT 15
A42214
peroxisome proliferator-activated receptor alpha chain - African clawed frog
N:Alternate names: PPAR-alpha
C:Species: Xenopus laevis (African clawed frog)
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 20-Sep-1999
A:Reference number: A42214; MUID:92191267; PMID:1312391
A:Accession: A42214
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-474 <DR>
A:Cross-references: GB:M84161; NID:9214663; PIDN:AAA49935.1; PID:9214664
A>Note: sequence extracted from NCBI backbone (NCBIR:88292)
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2003, 03:57:36 ; Search time 19 Seconds
(without alignments)
1125.054 Million cell updates/sec

Title: US-09-931-007A-1
Perfect score: 3508
Sequence: 1 MSETLGDSPIDSDSFDTDT.....KTEIDMSLPLQLQETIKDLY 688

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 156504 seqs, 31069816 residues
Total number of hits satisfying chosen parameters: 156504

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications-AA: *
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3508	100.0	688	US-09-931-007A-1	Sequence 1, Appl1
2	2231	63.6	505	US-09-765-111A-16	Sequence 16, Appl1
3	2196.5	60.6	506	US-10-109-886-6	Sequence 6, Appl1
4	2106.5	62.0	811	US-09-765-111A-23	Sequence 23, Appl1
5	2106.5	60.0	874	US-09-765-111A-6	Sequence 6, Appl1
6	2099	59.8	777	US-09-765-111A-2	Sequence 2, Appl1
7	2094	59.7	840	US-09-765-111A-4	Sequence 4, Appl1
8	2054.5	58.6	478	US-09-765-111A-27	Sequence 27, Appl1
9	2054	58.6	475	US-10-142-373-2	Sequence 2, Appl1
10	1076.5	30.7	440	US-10-013-807-2	Sequence 2, Appl1
11	524	14.9	121	US-09-925-299-1029	Sequence 1029, App
12	512.5	14.6	128	US-09-925-299-882	Sequence 882, App
13	392	11.2	1394	US-10-108-605-213	Sequence 213, App
14	385.5	11.0	1237	US-10-108-605-211	Sequence 211, App
15	333	9.5	81	US-09-864-761-35611	Sequence 35611, A
16	325.5	9.3	448	US-09-814-604-2	Sequence 2, Appl1
17	325.5	9.3	448	US-09-797-727-4	Sequence 4, Appl1
18	315	9.0	462	US-09-797-727-3	Sequence 3, Appl1
19	313	8.9	454	US-09-814-604-3	Sequence 3, Appl1

20	313	8.9	454	10	US-09-797-727-2	Sequence 2, Appl1
21	312	8.9	462	9	US-09-814-604-1	Sequence 1, Appl1
22	296	8.4	516	10	US-09-895-840-2	Sequence 2, Appl1
23	282.5	8.1	469	9	US-10-155-379-2	Sequence 2, Appl1
24	281.5	8.0	476	9	US-10-188-721-1	Sequence 1, Appl1
25	276.5	7.9	446	10	US-09-909-446-2	Sequence 2, Appl1
26	276.5	7.9	446	10	US-09-909-325-2	Sequence 2, Appl1
27	276.5	7.9	446	10	US-09-909-326-2	Sequence 2, Appl1
28	273.5	7.8	446	12	US-10-013-823-3	Sequence 2, Appl1
29	270.5	7.7	461	12	US-10-013-823-3	Sequence 3, Appl1
30	266	7.6	577	10	US-09-853-366-42	Sequence 42, Appl1
31	266	7.6	583	10	US-09-965-703-17	Sequence 17, Appl1
32	266	7.6	625	10	US-09-965-703-16	Sequence 16, Appl1
33	266	7.6	746	9	US-09-042-488B-5	Sequence 5, Appl1
34	266	7.6	746	9	US-09-042-488B-7	Sequence 7, Appl1
35	266	7.6	1041	9	US-09-042-488B-9	Sequence 9, Appl1
36	265	7.6	467	10	US-09-965-703-63	Sequence 63, Appl1
37	260.5	7.4	328	10	US-09-965-703-36	Sequence 36, Appl1
38	259.5	7.4	328	10	US-09-965-703-31	Sequence 31, Appl1
39	255.5	7.3	412	10	US-09-965-703-11	Sequence 11, Appl1
40	255.5	7.3	412	10	US-09-965-703-12	Sequence 12, Appl1
41	255.5	7.3	440	10	US-09-965-703-71	Sequence 71, Appl1
42	255.5	7.3	513	10	US-09-965-703-59	Sequence 59, Appl1
43	252.5	7.2	606	10	US-09-952-559-3	Sequence 3, Appl1
44	233.5	6.7	434	9	US-09-143-828-2	Sequence 2, Appl1
45	233.5	6.7	473	9	US-09-143-828-4	Sequence 4, Appl1

ALIGNMENTS

RESULT 1
US-09-931-007A-1
Sequence 1, Application US/09931007A
Patent No. US2002016132A1
GENERAL INFORMATION:
APPLICANT: Aventis Pharma S.A.
TITLE OF INVENTION: SYSTEM FOR REGULATING IN VIVO THE EXPRESSION OF A TRANSGENE BY
FILE OF INVENTION: CONDITIONAL INHIBITION
FILE REFERENCE: 03806.0512
CURRENT APPLICATION NUMBER: US/09/931,007A
CURRENT FILING DATE: 2001-08-17
PRIOR APPLICATION NUMBER: FR 00/10730
PRIOR FILING DATE: 2000-08-18
PRIOR APPLICATION NUMBER: US 60/239,246
PRIOR FILING DATE: 2000-10-11
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.0
SEQ ID NO 1
LENGTH: 688
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc
LOCATION: (1)..(688)
OTHER INFORMATION: Sequence for PPAR-gamma-2, a modified human PPAR-gamma
US-09-931-007A-1
Query Match
Best Local Similarity 100.0%; Score 3508; DB 9; Length 688;
Matches 688; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSETLGDSPIDSDSFDTDTLSANISQEWTVNDEMPFPTNFGISSVDLSWEDHSF 60
DB 1 MSETLGDSPIDSDSFDTDTLSANISQEWTVNDEMPFPTNFGISSVDLSWEDHSF 60
QY 61 DKPFTTVFSSISPHYEDIFETRPVADYKVDLKQEQSAIKVPPASPYSEKT 120
DB 61 DKPFTTVFSSISPHYEDIFETRPVADYKVDLKQEQSAIKVPPASPYSEKT 120
QY 121 QLYNRKCYCFQKCLAVGSMNAIRFGMPQAEKELAISSDIDQINDESADLRAL 180
DB 121 QLYNRKCYCFQKCLAVGSMNAIRFGMPQAEKELAISSDIDQINDESADLRAL 180

Db 121 QLYNRNKCQYCRFOKCLAVGMSHNAIRFGMRPOAEKEKLLAEISSDIDQNPESADLRAL 180
QY 181 AKHLUDYSIKSPPLTKAKARAILGKTTDKSPFIYDMNSLMGDEKIKRKHITPLQEOS 240
Db 181 AKHLUDYSIKSPPLTKAKARAILGKTTDKSPFIYDMNSLMGDEKIKRKHITPLQEOS 240
QY 241 KEVAIRIFOGCOPFRSVEAVOEITEYAKSIPGFVNLNDQVTLTKYGVHEIITYMLASLM 300
Db 241 KEVAIRIFOGCOPFRSVEAVOEITEYAKSIPGFVNLNDQVTLTKYGVHEIITYMLASLM 300
QY 301 NKQGVILISEGOGFMTREFLSLRKPFGEFMEPKFEFAVKFNALDELDDSDLAIFAVIILS 360
Db 301 NKQGVILISEGOGFMTREFLSLRKPFGEFMEPKFEFAVKFNALDELDDSDLAIFAVIILS 360
QY 361 GDRPGLNVPRIEDIONLQALELQKLNPRESSQLFAKLQKMTDLRQIVTEHVOLQ 420
Db 361 GDRPGLNVPRIEDIONLQALELQKLNPRESSQLFAKLQKMTDLRQIVTEHVOLQ 420
QY 421 VIKKTETMSLHPLLOEIRYKDLVAMAILTGKTTDKSPFIYDMNSLMGDEKIKRKHITP 480
Db 421 VIKKTETMSLHPLLOEIRYKDLVAMAILTGKTTDKSPFIYDMNSLMGDEKIKRKHITP 480
QY 481 LQDQSKVAIRIFOGCOPFRSVEAVOEITEYAKSIPGFVNLNDQVTLTKYGVHEIITYM 540
Db 481 LQDQSKVAIRIFOGCOPFRSVEAVOEITEYAKSIPGFVNLNDQVTLTKYGVHEIITYM 540
QY 541 LASLMNKDGLVISEGOGFMTREFLSLRKPFGEFMEPKFEFAVKFNALDELDDSDLAIFIA 600
Db 541 LASLMNKDGLVISEGOGFMTREFLSLRKPFGEFMEPKFEFAVKFNALDELDDSDLAIFIA 600
QY 601 VILSGDRPGLNVPRIEDIONLQALELQKLNPRESSQLFAKLQKMTDLRQIVTEH 660
Db 601 VILSGDRPGLNVPRIEDIONLQALELQKLNPRESSQLFAKLQKMTDLRQIVTEH 660
QY 661 VOLLOVYIKKTETMSLHPLLOEIRYKDL 688
Db 661 VOLLOVYIKKTETMSLHPLLOEIRYKDL 688

RESULT 2
US-09-765-111A-16
; Sequence 16, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-POLYMERIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-765-111A-16

Query Match 63.6%; Score 2231; DB 10; Length 505;
Best Local Similarity 87.7%; Pred. No. 8.6e-154;
Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 1 MGETLGDSPIDPESDSTDTLNSANISOEMTMVDEMPFMTNFGISSVDLSVMEHSHSF 60
Db 1 MGETLGDSPIDPESDSTDTLNSANISOEMTMVDEMPFMTNFGISSVDLSVMEHSHSF 60
QY 61 DIRPFTTVDFSSISTPHYEDIPFTRTDPVAVADKYKDLKLOEYOSAIKVEBPASPPYSEKT 120
Db 61 DIRPFTTVDFSSISTPHYEDIPFTRTDPVAVADKYKDLKLOEYOSAIKVEBPASPPYSEKT 120

Db 61 DIRPFTTVDFSSISTPHYEDIPFTRTDPVAVADKYKDLKLOEYOSAIKVEBPASPPYSEKT 120
QY 121 QLYN----- 124
Db 121 QLYNKPHEEPSNLSMAIECVCGDKASGFHYGVHACBGCGFFRRIRIRLKIYDRCDLNC 180
QY 125 -----RNKCOYCRFOKCLAVGMSHNAIRFGMRPOAEKEKLLAEISSDIDQNPESADLR 178
Db 181 RHNKSKNKCQYCRFOKCLAVGMSHNAIRFGMRPOAEKEKLLAEISSDIDQNPESADLR 240
QY 179 ALAKHLUDYSIKSPPLTKAKARAILGKTTDKSPFIYDMNSLMGDEKIKRKHITPLQ 238
Db 241 ALAKHLUDYSIKSPPLTKAKARAILGKTTDKSPFIYDMNSLMGDEKIKRKHITPLQ 300
QY 239 QSKVEAIRIFOGCOPFRSVEAVOEITEYAKSIPGFVNLNDQVTLTKYGVHEIITYMLAS 298
Db 301 QSKVEAIRIFOGCOPFRSVEAVOEITEYAKSIPGFVNLNDQVTLTKYGVHEIITYMLAS 360
QY 299 LMKNDGLVISEGOGFMTREFLSLRKPFGEFMEPKFEFAVKFNALDELDDSDLAIFAVIIL 358
Db 361 LMKNDGLVISEGOGFMTREFLSLRKPFGEFMEPKFEFAVKFNALDELDDSDLAIFAVIIL 420
QY 359 LSGDRPGLNVPRIEDIONLQALELQKLNPRESSQLFAKLQKMTDLRQIVTEHVOL 418
Db 421 LSGDRPGLNVPRIEDIONLQALELQKLNPRESSQLFAKLQKMTDLRQIVTEHVOL 480
QY 419 LOYIKKTETMSLHPLLOEIRYKDL 443
Db 481 LOYIKKTETMSLHPLLOEIRYKDL 505

RESULT 3
US-10-109-886-6
; Sequence 6, Application US/10109886
; Patent No. US20020119499A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO. LTD.
; APPLICANT: TANIGUCHI, Tomoyasu
; APPLICANT: MIZUKAMI, Junko
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND
; FILE REFERENCE: TANIGUCHI-6
; CURRENT APPLICATION NUMBER: US/10/109,886
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/514,247
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-109-886-6

Query Match 62.6%; Score 2196.5; DB 12; Length 506;
Best Local Similarity 86.8%; Pred. No. 2.7e-151;
Matches 439; Conservative 2; Mismatches 2; Indels 63; Gaps 2;

QY 1 MGETLGDSPIDPESDSTDTLNSANISOEMTMVDEMPFMTNFGISSVDLSVMEHSHSF 60
Db 1 MGETLGDSPIDPESDSTDTLNSANISOEMTMVDEMPFMTNFGISSVDLSVMEHSHSF 60
QY 61 DIRPFTTVDFSSISTPHYEDIPFTRTDPVAVADKYKDLKLOEYOSAIKVEBPASPPYSEKT 120
Db 61 DIRPFTTVDFSSISTPHYEDIPFTRTDPVAVADKYKDLKLOEYOSAIKVEBPASPPYSEKT 120
QY 121 QLYN----- 124
Db 121 QLYNKPHEEPSNLSMAIECVCGDKASGFHYGVHACBGCGFFRRIRIRLKIYDRCDLNC 180

Qy	125	-----RKQCYCFQKSLAVGMSHNAIRGRYPOAEKEVLAISSDIDQLNPESADR	178
Db	181	RHKKSRRKCCQCFQKSLAVGMSHNAIRGRYPOAEKEVLAISSDIDQLNPESADR	240
Qy	179	-ALAKHLDVYIKSPYLTKAKARAILGKTTDKSPFIYIMNSLMGEXIKRKHITPLQ	237
Db	241	QALAHNLVDYIKSPYLTKAKARAILGKTTDKSPFIYIMNSLMGEXIKRKHITPLQ	300
Qy	238	EOSKEVARIPOGOFNSVEVAOITEEAKSIPFVNLINDOVTLTKYGVHEIITMLA	297
Db	301	EOSKEVARIPOGOFNSVEVAOITEEAKSIPFVNLINDOVTLTKYGVHEIITMLA	360
Qy	298	SLMNKDGVLISEGQFMTREFLSLRKPFQGFMEKPEFAVKFNALELDDSDLAIFAVI	357
Db	361	SLMNKDGVLISEGQGFMTREFLSLRKPFQGFMEKPEFAVKFNALELDDSDLAIFAVI	420
Qy	368	ILSGDRPELWVKPIEDIQDNLOALDELQKLNPRESSQLPAKLQKMTDLROIVTEHVO	417
Db	421	ILSGDRPELWVKPIEDIQDNLOALDELQKLNPRESSQLPAKLQKMTDLROIVTEHVO	480
Qy	418	LLOYIKKETDMSLHPLLOEITKYOLY	443
Db	481	LLOYIKKETDMSLHPLLOEITKYOLY	506

```

RESULT 4
US-09-765-111A-23
Sequence 23, Application US/09765111A
Patent No. US20020106796A1
GENERAL INFORMATION:
APPLICANT: Fletcher, Jonathan A.
APPLICANT: Kroll, Todd G.
TITLE OF INVENTION: PAX8-PARAGANNA NUCLEIC ACID MOLECULES
TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
FILE REFERENCE: B0801/7196/ERP/MAT
CURRENT APPLICATION NUMBER: US/09/765,111A
CURRENT FILING DATE: 2001-01-18
PRIOR APPLICATION NUMBER: US 60/177,109
PRIOR FILING DATE: 2000-01-20
PRIOR APPLICATION NUMBER: US 60/225,079
PRIOR FILING DATE: 2000-08-14
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 811
TYPE: PRT
ORGANISM: Homo Sapiens
US-09-765-111A-23

```

Query Match	60.0%	Score 2106.5	DB 10	Length 811
Best Local Similarity	83.7%	Pred. No. 1.6e-14		
Matches 425	Conservative	4	Mismatches 14	Indels 65
				Gaps 3

QY	236	LOEÖSKEVAIRIFGÖCCÖRFSRSEAVOEITEIRAKSIRGEVNLNDLNOVTLTKYGVHEIITYTM	29
Db	604	LOEÖSKEVAIRIFGÖCCÖRFSRSEAVOEITEIRAKSIRGEVNLNDLNOVTLTKYGVHEIITYTM	665
QY	296	LASLNKKGCVLISEGÖGEMTREFELKSLKPGDMEKPEFEBAVVFNALEDDSDLAIFIA	355
Db	664	LASLNKKGCVLISEGÖGEMTREFELKSLKPGDMEKPEFEBAVVFNALEDDSDLAIFIA	722
QY	356	VILISGDRPGLLNVPRIEDIQNLQALOELOKLYLNHPRESSOLFAKLLQKMTDLQÖVYENH	415
Db	724	VILISGDRPGLLNVPRIEDIQNLQALOELOKLYLNHPRESSOLFAKLLQKMTDLQÖVYENH	783
QY	416	VOLLOVYIKKTEDMSLNLHELDEITYUDLY 443	
Db	784	VOLLOVYIKKTEDMSLNLHELDEITYUDLY 811	

```

RESULT 5
US-09-765-111A-6
? Sequence 6, Application US/09765111A
? Patent No. US20020106796A1
? GENERAL INFORMATION:
? APPLICANT: Fletcher, Jonathan A.
? APPLICANT: Kroil, Todd G.
? TITLE OF INVENTION: PAX8-PARAGAMMA NUCLEIC ACID MOLECULES
? TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
? FILE REFERENCE: 080801/77196/ERP/MAT
? CURRENT APPLICATION NUMBER: US/09/765,111A
? CURRENT FILING DATE: 2001-01-18
? PRIOR APPLICATION NUMBER: US 60/117,109
? PRIOR FILING DATE: 2000-01-20
? PRIOR APPLICATION NUMBER: US 60/225,079
? PRIOR FILING DATE: 2000-08-14
? NUMBER OF SEQ ID NOS: 47
? SOFTWARE: FASTSEQ for Windows Version 3.0
? SEQ ID NO 6
? LENGTH: 874
? TYPE: PRF
? ORGANISM: Homo Sapiens
US-09-765-111A-6

```

Query Match	60.0%;	Score 2106.5;	DB 10;	Length 874;
Best Local Similarity	83.7%;	Pred. No. 1.8e-144;		
Matches 425;	Conservative 4;	Mismatches 14;	Indels 65;	Gaps 3;

```

OY      356  VILISGDRPGLINVPKPIEDIQNLQALELQKLNHPSSQFLAKLQKMTDLQIVTEH 415
Db      787  VILISGDRPGLINVPKPIEDIQNLQALELQKLNHPSSQFLAKLQKMTDLQIVTEH 846

OY      416  VQLLOVIRKKTETDMSLHPLLOEIRYDLY 443
Db      847  VQLLOVIRKKTETDMSLHPLLOEIRYDLY 874

RESULT 6
US-09-765-111A-2
: Sequence 2, Application US/09765111A
: Patent No. US20020106796A1
: GENERAL INFORMATION:
: APPLICANT: Fletcher, Jonathan A.
: APPLICANT: Kroll, Todd G.
: TITLE OF INVENTION: PAX8-PARGAMMA NUCLEIC ACID MOLECULES
: TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
: FILE REFERENCE: B0801/7196/ERP/NAT
: CURRENT APPLICATION NUMBER: US/09/765,111A
: CURRENT FILING DATE: 2001-01-18
: PRIOR APPLICATION NUMBER: US 60/177,109
: PRIOR FILING DATE: 2000-01-20
: PRIOR APPLICATION NUMBER: US 60/225,079
: PRIOR FILING DATE: 2000-08-14
: NUMBER OF SEQ ID NOS: 47
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 777
: TYPE: PRT
: ORGANISM: Homo Sapiens
: US-09-765-111A-2

```

Query Match	59.8%;	Score 2099;	DB 10;	Length 777;
Best Local Similarity	83.3%;	Pred. No. 5.3e-144;		
Matches 420;	Conservative	5;	Mismatches 13;	Indels 66; Gaps 2

QY	6	GDSPIDPDSPTDPLTUSAN----	ISOEMVTOTEMPFWNTNGISSVDLSWEDHSHSD	61
Db	274	GKATITPSTNTPLGRRLSHLTQITPVVAEKMVTTEMPFWNTNGISSVDLSWEDHSHSD		333
QY	62	IKPFTTVDFSSISTPHUEDIPFTRTDPVADKKYDLKLOEYOSAIKVEEPASPPYSEKTO		121
Db	334	IKPFTTVDFSSISTPHUEDIPFTRTDPVADKKYDLKLOEYOSAIKVEEPASPPYSEKTO		393
QY	122	LYN-----		124
Db	394	LYNRPHEEPSNLSMAIECRVCGDKASGFHYGVNAECGCKGFPRRTIRLKIYDRCDLNCR		453
QY	125	-----RNRCOYCRPOKCLAVGSMHNAIRFGMRPOKEKELAEISSDIDOLPESADLRA		179
Db	454	IHKTSRMCQYCRPOKCLAVGSMHNAIRFGMRPOKEKELAEISSDIDOLPESADLRA		513
QY	180	LAKHLUDYCIKSFPLTKAKARAILGKTITDKSPFIYDMNSLMGEDIKFKHITPPILOQ		239
Db	514	LAKHLUDYCIKSFPLTKAKARAILGKTITDKSPFIYDMNSLMGEDIKFKHITPPILOQ		573
QY	240	SKEVAIRIQGCGQFSRSEVAVOEITEYAKSIPGFVNLIDLNDQYTLKYGVAHEIITYMLASL		299
Db	574	SKEVAIRIQGCGQFSRSEVAVOEITEYAKSIPGFVNLIDLNDQYTLKYGVAHEIITYMLASL		633
QY	300	MNKDGVLISEGOGFRTREFLSLKRPFQGFMEPKREFPAVKFNALBIDSDLAIFIAVITL		359
Db	634	MNKDGVLISEGOGFRTREFLSLKRPFQGFMEPKREFPAVKFNALBIDSDLAIFIAVITL		693
QY	360	SGDRGLLNVRPIEDIONLLOALDLQKLKNRPESSQLFAKLLQKMTDRLQIVTEHYOLL		419
Db	694	SGDRGLLNVRPIEDIONLLOALDLQKLKNRPESSQLFAKLLQKMTDRLQIVTEHYOLL		753
QY	420	QVIKKTETDMSLHPLQETIYKDL 443		
Db	754	QVIKKTETDMSLHPLQETIYKDL 777		

```

RESULT 7
US-09-765-111A-4
; Sequence 4, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PARAGAMMA NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B080177196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 840
; TYPE: PrT
; ORGANISM: Homo Sapiens
; US-09-765-111A-4

```

Query Match	59.7%;	Score 2094;	DB 10;	Length 840;
Best Local Similarity	87.0%;	Pred. No. 1.3e-143;		
Matches 416; Conservative	0;	Mismatches 0;	Indels 62;	Gaps 1.

Qy	28	EMMVTDEMFMTNGISVSLVMEHSHSFOLKPTTYVDFSSITPHYEDLPFRTD	87
Db	363	EMMVTDEMFMTNGISVSLVMEHSHSFOLKPTTYVDFSSITPHYEDLPFRTD	422
Qy	88	PVVADYKYLKLOEYOSAIKVEPASPYYSEKTOLYN-----	124
Db	423	PVVADYKYLKLOEYOSAIKVEPASPYYSEKTOLYNPHBEPNSLMAIECRVCGKAS	482
Qy	125	-----RKOCYCRFOCLAVGSHNA	145
Db	483	GPHYGVHAGECCKGFERRTRILKLIYDRCLNCRHKSRKCOYCRFOCLAVGSHNA	542
Qy	146	IRGRMPQAEKEKLLAEISSDIQOLNPESADRLAKLKHLYSYIKSFPLRKAKARALITG	205
Db	543	IRGRMPQAEKEKLLAEISSDIQOLNPESADRLAKLKHLYSYIKSFPLRKAKARALITG	602
Qy	206	KTTDKSPFIYIYMNSLMGEBKIKRKHITPLQEOBSKEVALIFGCGCFRSEVAEOETTEX	265
Db	603	KTTDKSPFIYIYMNSLMGEBKIKRKHITPLQEOBSKEVALIFGCGCFRSEVAEOETTEX	662
Qy	266	AKSIPFVNLDLNDQVTLKLYGVHEIITYMLASLMNKGVLISGQGFMTREFLKSLRP	325
Db	663	AKSIPFVNLDLNDQVTLKLYGVHEIITYMLASLMNKGVLISGQGFMTREFLKSLRP	722
Qy	326	FGDFMEKFEPAVKNALIEDSDLAIFIAVILISGDRPGLLVKPIEDIQMLQALEL	385
Db	723	FGDFMEKFEPAVKNALIEDSDLAIFIAVILISGDRPGLLVKPIEDIQMLQALEL	782
Qy	386	QLKLNPESSQLEAKLKQMTDLROIVYEHQOLLQVIRKFTEDMSLPLQLOEYIKDLY	443
Db	783	QLKLNPESSQLEAKLKQMTDLROIVYEHQOLLQVIRKFTEDMSLPLQLOEYIKDLY	840

```

RESULT 8
US-09-765-111A-27
; Sequence 27, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PARAMMA NUCLEIC ACID MOLECULES
; AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/1196/ERP/MAT

```


Db 175 KAFSKHIYNALVKNFMNMTKKARSILTGKSSHNAPEVLIHDIETLMOAERGLWVKQVLNGL 234
QY 238 EQSEVAIRIFOGQGFSEVAEVOITEYAKSIPGFVNLNDVOYTLTKYGHETIYMLA 297
Db 235 PPIVEISVHFYRQSTVETVELTEFAKNIPNFSSLPDVOYTLTKYGHETIYMLA 294
QY 298 SLMNKDGVLISEGQGFMTREFLKSIRKPFQDFMPEKFEFAVFNALIEDSDLAIFIAVI 357
Db 295 SIYVKDGLLVANGSGFYTHEFLRSRKPFSDIIEPKFEFANFKFALIEDSDLAIFIAI 354
QY 358 ILSGDRPGLLVKPIEDIODNLQALQKLNHPSSQLEFAKLQKMTDIKQIVTEHVQ 417
Db 355 ILGDRPGLLVNVPQVYAIODTILFALFPHLQVNHPSQYLFPKLQKADLRQLVTEHAQ 414
QY 418 LLOVIRKTEMDMSLHPLLOETIKDLY 443
Db 415 MMOWLKTESETLHPLLOETIKDLY 440

RESULT 11

US-09-925-299-1029
; Sequence 1029, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1029
; LENGTH: 121
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (9)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; LOCATION: (108)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-1029

Query Match 14.9%; Score 524; DB 10; Length 121;
Best Local Similarity 91.4%; Pred. No. 2,8e-31;
Matches 106; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 462 DMSLMGGEKIRFKHTITPLQEOSKEVAIRIFOGQGFSEVAEVOITEYAKSIPGFVNL 521
Db 1 DMSLMGGEKIRFKHTITPLQEOSKEVAIRIFOGQGFSEVAEVOITEYAKSIPGFVNL 60
QY 522 LNDVOYTLTKYGHETIYMLASLMNKDGLVISEGQGFMTREFLKSIRKPFQDFNEP 577
Db 61 LNDVOYTLTKYGHETIYMLASLMNKDGLVISEGQGFMTREFLKSIRKPFQDFNEP 116

RESULT 12

US-09-925-299-882
; Sequence 882, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 882
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (112)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-299-882

Query Match 14.6%; Score 512.5; DB 10; Length 128;
Best Local Similarity 85.2%; Pred. No. 2e-30;
Matches 98; Conservative 2; Mismatches 4; Indels 11; Gaps 1;

QY 22 SANISQ-----EMTWDTMPFMPNFGISSVDLSVMDHSHSFDIKPFTTVDF 70
Db 1 TTNIQGHFLKRESAFNEMTMVDTMPFTNFGISSVDLSVMDHSHSFDIKPFTTVDF 60
QY 71 SSISTPHYEDIPFTRPDVADKYDKLQEOYSAIYVEPASPYYSEKQOLYNR 125
Db 61 SSISTPHYEDIPFTRPDVADKYDKLQEOYSAIYVEPASPYYSEKQOLYNR 115

RESULT 13

US-10-108-605-213
; Sequence 213, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCOD
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 213
; LENGTH: 1394
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-213

Query Match 11.2%; Score 392; DB 9; Length 1394;
Best Local Similarity 27.8%; Pred. No. 2,6e-20;
Matches 121; Conservative 80; Mismatches 160; Indels 74; Gaps 16;

QY 48 VDLVMEHSHSFDIKPFTTVDFSSISTPHYEDIPFTRPVV-----ADYKDLKQ-- 100
Db 382 LDSSQLVYLCQKDFDEKIDTALSNSSANT-----GRNTPAVTANEDADGFRRSIQOK 433
QY 101 -EYOSAIKVEPASPYYSEKQOLYNRKCOYCFORCLAVGSMHAIREGRMPOAKEKL 159
Db 434 IYRPTCKNQC-----ILRINRRCQYCRKLKCIAGMSRDVAFREGVAKREKARI 486
QY 160 LAETSSIIDLNEBSADRLAKHLVD-----SYKSPPLTKAKARALLTGKTTD 209
Db 487 LAA-----MOOSTONRGQORALATELDQPRLLAVALRAHLETEFTRKEKVSARORARD 542
QY 210 KSPFVIYDMNSLMGGEKIRFKHTITPLQEOSKEVAIRIFOGQGFSEVAEVOITEYAKS 269

```
Db 543 PS-----YSMPTLLACP-----LNPAPELQSE-----OEFSGREAHVINGVIDFAGMI 585
OY 270 PGFVNILDLNOQVTLTKGYVEHIIYTMLASIMNK--DGVLLSEGGFMTREFLKS-LKRPF 326
Db 586 PGFOLLQDDDKFTLLKAGLDALFVRLICMFDDSSINIIICLNG-VVRRAIOLGANARF 644
OY 327 GDFMEPFEEFAVFNALDSDLAIFIAVIIISGDRPGLLNVPEDIODNLQALELQ 386
Db 645 --LVDSTFNFAERNNSNNLDAELGFCALVILTPDRPGRLNELLEKMTSRKGCIOYI 702
OY 387 LKLNHPSSQLFAKLQKMTDLROIVTEHVQLLQVYIKKTTSDMLHPLOEIKDYLA 446
Db 703 VAONRRPQPEFLAKLLETMTDRLTLSTLHTEKLVFTEHEKEL-----LROOM-----WS 752
OY 447 ILGKTTD---KSP 457
Db 753 MEDGNSDGOQNKSP 767

RESULT 14
US-10-108-605-211
; Sequence 211, Application US/10108605
; Patent No. US20020160934A1
; GENERAL INFORMATION:
; APPLICANT: Broadus, Julie
; APPLICANT: Stam, Lynn
; APPLICANT: Bachmann, Jane
; APPLICANT: Kamdar, Kim
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
; FILE REFERENCE: 31133B
; CURRENT APPLICATION NUMBER: US/10/108,605
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: US 09/761,142
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/176,418
; PRIOR FILING DATE: 2000-01-14
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 211
; LENGTH: 1237
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-10-108-605-211

Query Match 11.0%; Score 385.5; DB 9; Length 1237;
Best Local Similarity 29.3%; Pred. No. 6.4e-20;
Matches 103; Conservative 71; Mismatches 126; Indels 51; Gaps 12;

OY 124 NRKRCQYCFQKCLAVGMSHNAIRFGMPQAEKEKLLAEISSIDOLNPESADLRALAKH 183
Db 294 NRKRCQYCRKCKICINAVGMSDAVRFQGVPRKREKARILAA-----MOOSTORGOORALATE 349
OY 184 LVD-----SYTKSPETYAKARAILTGTTDKSPFVIYDMSLMGDEKIRFKHI 233
Db 350 LDDQPRLLAAVLAHLHETCEFTKEKVSAMRQARDCPS-----YSMPTLLACP-----L 398
OY 234 TPLQEQSKFAIRIFQCGCFRSEVAOETETVAKSTIPGFVNLDLNOQVTLTKGYVEHIIY 293
Db 399 NPPELQSE-----OEFSGREAHVINGVIDFAGMIPEGOLLQDDKFTLLKAGLDALF 452
OY 294 TMLASIMNK--DGVLLSEGGFMTREFLKS-LKRPFGDFMEPFEEFAVFNALDSDLDL 350
Db 453 VRLICMFDDSSINIIICLNG-VVRRAIOLGANARF--LVDSTFNFAERNNSNNLDAEL 509
OY 351 AITFAVITISGDRPGLLNKVPEDIODNLQALELQKLNHPSSQLFAKLQKMTDLRO 410
Db 510 GLFCALVILTPDRPGRLNELLEKMTSRKGCIOYIQAONRRPQPEFLAKLLETMTDRLR 569
OY 411 IYVEHQVOLLQVYIKKTTSDMLHPLOEIKDYLAAILGKTTD---KSP 457
Db 570 LSTLHTEKLVFTEHEKEL-----LROOM-----WSMEDGNSDGOQNKSP 610
```

```
RESULT 15
US-09-864-761-35611
; Sequence 35611, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35611
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO NM_007104.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4
; OTHER INFORMATION: EST_HUMAN HIT: AUI31142.1, EVALUE 5.00e-39
; OTHER INFORMATION: SWISSPROT HIT: Q03181, EVALUE 6.00e-40
US-09-864-761-35611
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 25, 2003, 03:58:31 ; Search time 22 Seconds
(without alignments)
920.135 Million cell updates/sec

Title: US-09-931-007A-1
Perfect score: 3508
Sequence: 1 MGETLGDSPIDESDSFTDT.....KETDMSLHPLLQEIYKDLX 688

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2231	63.6	505	4	US-09-128-142-4 Sequence 4, Appli
2	2196.5	62.6	506	4	US-09-514-247A-6 Sequence 6, Appli
3	2089	59.5	477	4	US-09-128-142-2 Sequence 2, Appli
4	2054	58.6	475	2	US-08-484-200-2 Sequence 2, Appli
5	2054	58.6	475	3	US-08-465-375-2 Sequence 2, Appli
6	2054	58.6	475	4	US-09-788-070-2 Sequence 2, Appli
7	2044.5	58.3	476	4	US-08-134-557D-2 Sequence 2, Appli
8	2043	58.2	475	4	US-08-764-870-10 Sequence 10, Appli
9	2043	58.2	475	4	US-08-980-115-9 Sequence 10, Appli
10	1101.5	31.4	441	4	US-08-764-870-9 Sequence 9, Appli
11	1101.5	31.4	441	4	US-08-980-115-9 Sequence 9, Appli
12	1081.5	30.8	468	1	US-08-459-287-3 Sequence 4, Appli
13	1076.5	30.7	440	2	US-08-484-200-4 Sequence 4, Appli
14	1076.5	30.7	468	1	US-08-459-287-2 Sequence 2, Appli
15	1074.5	30.6	468	4	US-08-764-870-8 Sequence 8, Appli
16	1074.5	30.6	468	4	US-08-980-115-8 Sequence 8, Appli
17	1071	30.5	439	1	US-08-333-358-10 Sequence 10, Appli
18	1071	30.5	439	1	US-08-463-694-10 Sequence 10, Appli
19	1071	30.5	439	1	US-08-694-501-10 Sequence 10, Appli
20	375.5	10.7	483	4	US-09-106-194-4 Sequence 4, Appli
21	341.5	9.7	548	1	US-08-333-358-2 Sequence 2, Appli
22	341.5	9.7	548	1	US-08-463-694-2 Sequence 2, Appli
23	341.5	9.7	548	1	US-08-694-501-2 Sequence 2, Appli
24	341.5	9.7	556	4	US-09-106-194-3 Sequence 3, Appli
25	324.5	9.3	368	6	5223606-3 Patent No. 5223606
26	324.5	9.3	448	6	5223606-2 Patent No. 5223606
27	317	9.0	416	4	US-08-764-870-4 Sequence 4, Appli

28	317	9.0	416	4	US-08-980-115-4 Sequence 4, Appli
29	316	9.0	462	6	5171671-2 Patent No. 5171671
30	315	9.0	462	2	US-08-592-383-2 Sequence 2, Appli
31	315	9.0	462	2	US-08-095-728B-4 Sequence 4, Appli
32	315	9.0	462	5	PCT-US92-02320A-4 Sequence 2, Appli
33	315	9.0	462	5	PCT-US92-02320A-4 Sequence 2, Appli
34	315	9.0	797	5	US-08-095-728B-2 Sequence 2, Appli
35	313	8.9	454	4	PCR-US92-02320A-2 Sequence 2, Appli
36	313	8.9	454	4	US-08-764-870-5 Sequence 5, Appli
37	313	8.9	454	4	US-08-980-115-5 Sequence 5, Appli
38	305	8.7	440	1	US-08-333-358-8 Patent No. 5260432
39	305	8.7	440	1	US-08-463-694-8 Sequence 8, Appli
40	305	8.7	440	1	US-08-694-501-8 Sequence 8, Appli
41	305	8.7	447	1	US-08-373-935-1 Sequence 1, Appli
42	296	8.4	557	4	US-08-653-648A-12 Sequence 12, Appli
43	293.5	8.4	403	2	US-08-592-383-4 Sequence 4, Appli
44	283.5	8.1	674	4	US-08-653-648A-14 Sequence 14, Appli
45	283	8.1	433	2	US-08-466-120-2 Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-128-142-4
; Sequence 4, Application US/09128142
; Patent No. 6284559
GENERAL INFORMATION:
APPLICANT: Smith, Roy G.
TITLE OF INVENTION: ANTIPROLIFERATIVE AGENTS ASSOCIATED WITH PEROXISOME PROLIFERATOR ACTIVATED RECEPTORS GAMMA1 AND G
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Jack L. Tribble
STREET: 126 E. Lincoln Ave., P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Power Macintosh 7500/100
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/128,142
FILING DATE: 03-Aug-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/844,007
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Tribble, Jack L.
REGISTRATION NUMBER: 32,633
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-5321
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-128-142-4
Query Match 63.6%; Score 2231; DB 4; Length 505;
Best Local Similarity 87.7%; Pred. No. 7.6e-186;
Matches 443; Conservative 0; Mismatches 62; Gaps 1;
QY 1 MGETLGDSPIDESDSFTDTLSANISQEMTWVDTMFPWPTNFGISSVSLVMEHSHSF 60

```
|||||
Db 1 MGETLDSPIIDPSDSFTDTLSANISQEMTMVDEMPFWMTNNGISSVDLSVMEDSHSF 60
QY 61 DLRPFTTVDFSSISTPHYEDIPFTRTDPVADYKYDKLOEYOSAIKVEPASPYYSEKT 120
Db 61 DLRPFTTVDFSSISTPHYEDIPFTRTDPVADYKYDKLOEYOSAIKVEPASPYYSEKT 120
QY 121 QLYN----- 124
Db 121 QLYNKHHEEPSNSLMAIECRVCGDKASGFHYGVHACGCKGFFRRITRLKLIYDRCDLNC 180
QY 125 -----RNCQYCRFOKCLAVGSHNAIRGRMPQAEKEKLAIESSDIDQLNPESADLR 178
Db 181 RIHKSRNKCQYCRFOKCLAVGSHNAIRGRMPQAEKEKLAIESSDIDQLNPESADLR 240
QY 179 ALAKHLYDSYIKSFPLTKAKARAILTGKTTDKSPVIYDMNSLMMGEDKIKFKHITPLOE 238
Db 241 ALAKHLYDSYIKSFPLTKAKARAILTGKTTDKSPVIYDMNSLMMGEDKIKFKHITPLOE 300
QY 239 OSKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLDLNDQVTLTKYGVHEIITYMLAS 298
Db 301 OSKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLDLNDQVTLTKYGVHEIITYMLAS 360
QY 299 LANKDGVLSIEGQGFMTREFLKSRLKPPGDFEMPEKFEFAVKFNALDELDDSLAIFIAVII 358
Db 361 LANKDGVLSIEGQGFMTREFLKSRLKPPGDFEMPEKFEFAVKFNALDELDDSLAIFIAVII 420
QY 359 LSGDRGGLNVKPIEDIDQNLQALELOLKLNPRESSQLFAKLQKMTDLROIVTEHVQ 418
Db 421 LSGDRGGLNVKPIEDIDQNLQALELOLKLNPRESSQLFAKLQKMTDLROIVTEHVQ 480
QY 419 LQVIKKTEMDMSLHPLOEITYKDL 443
Db 481 LQVIKKTEMDMSLHPLOEITYKDL 505
```

RESULT 2

```
US-09-514-247A-6
; Sequence 6, Application US/09514247A
; Patent No. 6365361
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO. LTD.
; APPLICANT: TANIGUCHI, Tomoyasu
; APPLICANT: MIZUKAMI, Junko
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPA
; FILE REFERENCE: TANIGUCHI-6
; CURRENT APPLICATION NUMBER: US/09/514,247A
; CURRENT FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-514-247A-6
```

Query Match 62.6%; Score 2196.5; DB 4; Length 506;
Best Local Similarity 86.8%; Pred. No. 7.6e-183;
Matches 439; Conservative 2; Mismatches 2; Indels 63; Gaps 2;

```
QY 1 MGETLDSPIIDPSDSFTDTLSANISQEMTMVDEMPFWMTNNGISSVDLSVMEDSHSF 60
Db 1 MGETLDSPIIDPSDSFTDTLSANISQEMTMVDEMPFWMTNNGISSVDLSVMEDSHSF 60
QY 61 DLRPFTTVDFSSISTPHYEDIPFTRTDPVADYKYDKLOEYOSAIKVEPASPYYSEKT 120
Db 61 DLRPFTTVDFSSISTPHYEDIPFTRTDPVADYKYDKLOEYOSAIKVEPASPYYSEKT 120
QY 121 QLYN----- 124
```

```
|||||
Db 121 QLYNKHHEEPSNSLMAIECRVCGDKASGFHYGVHACGCKGFFRRITRLKLIYDRCDLNC 180
QY 125 -----RNCQYCRFOKCLAVGSHNAIRGRMPQAEKEKLAIESSDIDQLNPESADLR 178
Db 181 RIHKSRNKCQYCRFOKCLAVGSHNAIRGRMPQAEKEKLAIESSDIDQLNPESADLR 240
QY 179 -ALAKHLYDSYIKSFPLTKAKARAILTGKTTDKSPVIYDMNSLMMGEDKIKFKHITPLO 237
Db 241 QALAKHLYDSYIKSFPLTKAKARAILTGKTTDKSPVIYDMNSLMMGEDKIKFKHITPLO 300
QY 238 OSKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLDLNDQVTLTKYGVHEIITYMLA 297
Db 301 OSKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLDLNDQVTLTKYGVHEIITYMLA 360
QY 298 SLNKKDGVLSIEGQGFMTREFLKSRLKPPGDFEMPEKFEFAVKFNALDELDDSLAIFIAVI 357
Db 361 SLNKKDGVLSIEGQGFMTREFLKSRLKPPGDFEMPEKFEFAVKFNALDELDDSLAIFIAVI 420
QY 358 ILSGDRGGLNVKPIEDIDQNLQALELOLKLNPRESSQLFAKLQKMTDLROIVTEHVQ 417
Db 421 ILSGDRGGLNVKPIEDIDQNLQALELOLKLNPRESSQLFAKLQKMTDLROIVTEHVQ 480
QY 418 LQVIKKTEMDMSLHPLOEITYKDL 443
Db 481 LQVIKKTEMDMSLHPLOEITYKDL 506
```

RESULT 3

```
US-09-128-142-2
; Sequence 2, Application US/09128142
; Patent No. 6294559
; GENERAL INFORMATION:
; APPLICANT: Smith, Roy G.
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENTS ASSOCIATED WITH
; PEROXISOME PROLIFERATOR ACTIVATED RECEPTORS GAMMA1 AND C
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Jack L. Tribble
; STREET: 126 E. Lincoln Ave., P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Power Macintosh 7500/100
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,142
; FILING DATE: 03-Aug-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,007
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Tribble, Jack L.
; REGISTRATION NUMBER: 32,633
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-5321
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-128-142-2
```

```

Query Match          59.5%; Score 2089; DB 4; Length 477;
Best Local Similarity 87.0%; Pred. No. 1.5e-173;
Matches 415; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 29 MTDVTEPMPNFGISSVDLSVMDHSHSFDIKPTTVDFSSISTPHYEDIPPTPTDP 88
DB 1 MTDVTEPMPNFGISSVDLSVMDHSHSFDIKPTTVDFSSISTPHYEDIPPTPTDP 60

QY 89 VVADYKDKLQEOYSAIKVEPASPYYSEKTOLYN-----RNKCQYCRFOKCLAVGSHNAI 124
DB 61 VVADYKDKLQEOYSAIKVEPASPYYSEKTOLYNKRPHPEPSNSLMAIECRVCGDKASG 120

QY 125 -----RNKCQYCRFOKCLAVGSHNAI 146
DB 121 FHYGVHAGECGKGFERRITRLKLIYDRCDLNCRIHKSRNKCQYCRFOKCLAVGSHNAI 180

QY 147 RFGMPQAEKEXLLAEISSDIDOLNPESADLRALAKHLVDSYIKSPPLTKAKARAILTGK 206
DB 181 RFGMPQAEKEXLLAEISSDIDOLNPESADLRALAKHLVDSYIKSPPLTKAKARAILTGK 240

QY 207 TTDKSPFVIYDMNSLMMGEDKIKFKHITPLOEQSKEVAIRIFOGCQFSPSYEAVOEITEYA 266
DB 241 TTDKSPFVIYDMNSLMMGEDKIKFKHITPLOEQSKEVAIRIFOGCQFSPSYEAVOEITEYA 300

QY 267 KSIIPGEVNDLNDQVTLKYGVEHIIYTMLASLMNKDGLVISEGQGFMTREFLKSRRPF 326
DB 301 KSIIPGEVNDLNDQVTLKYGVEHIIYTMLASLMNKDGLVISEGQGFMTREFLKSRRPF 360

QY 327 GDMPEKFEFAVKFNALIEDSDLAIFIAVIIISGDRGLLNKPIEDIODNLQALELQ 386
DB 361 GDMPEKFEFAVKFNALIEDSDLAIFIAVIIISGDRGLLNKPIEDIODNLQALELQ 420

QY 387 LKLNHPSSQLFAKLQKMTDLRQIVTEHVQLOLVIKKTEETDMSLHPLLOEITKDY 443
DB 421 LKLNHPSSQLFAKLQKMTDLRQIVTEHVQLOLVIKKTEETDMSLHPLLOEITKDY 477

RESULT 4
US-08-484-200-2
; Sequence 2, Application US/08484200
; Patent No. 5861274
; GENERAL INFORMATION:
; APPLICANT: EVANS, RONALD M.
; APPLICANT: FORMAN, BARRY M.
; APPLICANT: KLIEMER, STEVEN A.
; APPLICANT: ONG, ESTELITA S.
; TITLE OF INVENTION: NOVEL MEMBERS OF THE STEROID/THYROID
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: PRETTY, SCHROEDER, BRUEGEMANN & CLARK
; STREET: 444 SOUTH FLOWER STREET, SUITE 2000
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,200
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: REITER, STEPHEN E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9971
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-546-1995
; TELEFAX: 619-546-9392
; INFORMATION FOR SEQ ID NO: 2:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-484-200-2

Query Match          58.6%; Score 2054; DB 2; Length 475;
Best Local Similarity 85.5%; Pred. No. 1.7e-170;
Matches 406; Conservative 4; Mismatches 3; Indels 62; Gaps 1;

QY 31 MVDTEPMPNFGISSVDLSVMDHSHSFDIKPTTVDFSSISTPHYEDIPPTPTDPV 90
DB 1 MVDTEPMPNFGISSVDLSVMDHSHSFDIKPTTVDFSSISADHYEDIPPTRADPV 60

QY 91 ADYKDKLQEOYSAIKVEPASPYYSEKTOLYN-----RNKCQYCRFOKCLAVGSHNAIRF 124
DB 61 ADYKDKLQEOYSAIKVEPASPYYSEKTOLYNRRPHPEPSNSLMAIECRVCGDKASGFH 120

QY 125 -----RNKCQYCRFOKCLAVGSHNAIRF 148
DB 121 YGVHAGECGKGFERRITRLKLIYDRCDLNCRIHKSRNKCQYCRFOKCLAVGSHNAIRF 180

QY 149 GRMPQAEKEXLLAEISSDIDOLNPESADLRALAKHLVDSYIKSPPLTKAKARAILTGKT 208
DB 181 GRMPQAEKEXLLAEISSDIDOLNPESADLRALAKHLVDSYIKSPPLTKAKARAILTGKT 240

QY 209 DKSPFVIYDMNSLMMGEDKIKFKHITPLOEQSKEVAIRIFOGCQFSPSYEAVOEITEYAS 268
DB 241 DKSPFVIYDMNSLMMGEDKIKFKHITPLOEQSKEVAIRIFOGCQFSPSYEAVOEITEYAS 300

QY 269 IPGFVNDLNDQVTLKYGVEHIIYTMLASLMNKDGLVISEGQGFMTREFLKSRRPFG 328
DB 301 IPGFVNDLNDQVTLKYGVEHIIYTMLASLMNKDGLVISEGQGFMTREFLKSRRPFG 360

QY 329 FMEKFEFAVKFNALIEDSDLAIFIAVIIISGDRGLLNKPIEDIODNLQALELQ 388
DB 361 FMEKFEFAVKFNALIEDSDLAIFIAVIIISGDRGLLNKPIEDIODNLQALELQ 420

QY 389 LKLNHPSSQLFAKLQKMTDLRQIVTEHVQLOLVIKKTEETDMSLHPLLOEITKDY 443
DB 421 LKLNHPSSQLFAKLQKMTDLRQIVTEHVQLOLVIKKTEETDMSLHPLLOEITKDY 475

RESULT 5
US-08-465-375-2
; Sequence 2, Application US/08465375A
; Patent No. 6022897
; GENERAL INFORMATION:
; APPLICANT: Evans, Ronald M.
; APPLICANT: Forman, Barry M.
; TITLE OF INVENTION: SELECTIVE MODULATORS OF PEROXISOME
; TITLE OF INVENTION: PROLIFERATOR ACTIVATED RECEPTOR-GAMMA, AND METHODS FOR THE
; FILE REFERENCE: SALK1470-1
; CURRENT APPLICATION NUMBER: US/08/465,375A
; CURRENT FILING DATE: 1995-06-05
; EARLIER APPLICATION NUMBER: 08/428,559
; SOFTWARE: FastSRO for Windows Version 4.0
; SEQ ID NO: 2
; LENGTH: 475
; TYPE: prt
; ORGANISM: Mus musculus
US-08-465-375-2

Query Match          58.6%; Score 2054; DB 3; Length 475;
Best Local Similarity 85.5%; Pred. No. 1.7e-170;
Matches 406; Conservative 4; Mismatches 3; Indels 62; Gaps 1;

QY 31 MVDTEPMPNFGISSVDLSVMDHSHSFDIKPTTVDFSSISTPHYEDIPPTPTDPV 90

```

```

Db 1 MVDEMPFMTNFGISSVDLSVMEHSHSPDIKPFVTVDSSISAPHYEDIPFTRADPMV 60
QY 91 ADVKYDLKLOEYOSAIKVEBPASPPYSEKTOLYN-----124
Db 61 ADVKYDLKLOEYOSAIKVEBPASPPYSEKTOLYNRPHEEPSNSLMAIECRVCGDKASGFH 120
QY 125 -----RNKCQYCRFOKCLAVGSMHNAIRF 148
Db 121 YGVHACGCKGFFRRTIRLKLIVDRCDLNCRIHKSSRNKCQYCRFOKCLAVGSMHNAIRF 180
QY 149 GRMPOAKEKELIAEISSDIDOLNPESADLRALAKHLVDYSIKSPPLTKAKARAILTGKTT 208
Db 181 GRMPOAKEKELIAEISSDIDOLNPESADLRALAKHLVDYSIKSPPLTKAKARAILTGKTT 240
QY 209 DKSPFVIYDMNSLMMGDKIKFKHITPLOSKEVAIRIFOGCOFRSEVAEOETTEYAKS 268
Db 241 DKSPFVIYDMNSLMMGDKIKFKHITPLOSKEVAIRIFOGCOFRSEVAEOETTEYAKN 300
QY 269 IPGFVNLNDQVTLTKYGVHEIITMLASLMNKDGVLSISGOGFMTREFLSLRKPFCD 328
Db 301 IPGFVNLNDQVTLTKYGVHEIITMLASLMNKDGVLSISGOGFMTREFLSLRKPFCD 360
QY 329 FMEKFEFAKFNALIEDDSDLAIFIAVITISGDRPGLLVKPIEDIQDNLQALEYOLK 388
Db 361 FMEKFEFAKFNALIEDDSDLAIFIAVITISGDRPGLLVKPIEDIQDNLQALEYOLK 420
QY 389 LNHESQSLFAKLQKMTDLROIVTEHVOLLQVITKKTETMSLHPLLOEYIKDLY 443
Db 421 LNHESQSLFAKLQKMTDLROIVTEHVOLLQVITKKTETMSLHPLLOEYIKDLY 475

```

```

RESULT 6
US-09-788-070-2
; Sequence 2, Application US/09788070
; Patent No. 6413994
; GENERAL INFORMATION:
; APPLICANT: EVANS, Ronald
; TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR ACTIVATED RECEPTOR-GAMMA,
; FILE REFERENCE: SALK1480-2
; CURRENT APPLICATION NUMBER: US/09/788, 070
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/955, 302
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Mus Musculus
US-09-788-070-2

```

```

Query Match 58.6%; Score 2054; DB 4; Length 475;
Best Local Similarity 85.5%; Pred. No. 1.7e-170;
Matches 406; Conservative 4; Mismatches 3; Indels 62; Gaps 1;

```

```

QY 31 MVDEMPFMTNFGISSVDLSVMEHSHSPDIKPFVTVDSSISAPHYEDIPFTRADPMV 90
Db 1 MVDEMPFMTNFGISSVDLSVMEHSHSPDIKPFVTVDSSISAPHYEDIPFTRADPMV 60
QY 91 ADVKYDLKLOEYOSAIKVEBPASPPYSEKTOLYN-----124
Db 61 ADVKYDLKLOEYOSAIKVEBPASPPYSEKTOLYNRPHEEPSNSLMAIECRVCGDKASGFH 120
QY 125 -----RNKCQYCRFOKCLAVGSMHNAIRF 148
Db 121 YGVHACGCKGFFRRTIRLKLIVDRCDLNCRIHKSSRNKCQYCRFOKCLAVGSMHNAIRF 180
QY 149 GRMPOAKEKELIAEISSDIDOLNPESADLRALAKHLVDYSIKSPPLTKAKARAILTGKTT 208
Db 181 GRMPOAKEKELIAEISSDIDOLNPESADLRALAKHLVDYSIKSPPLTKAKARAILTGKTT 240

```

```

QY 209 DKSPFVIYDMNSLMMGDKIKFKHITPLOSKEVAIRIFOGCOFRSEVAEOETTEYAKS 268
Db 241 DKSPFVIYDMNSLMMGDKIKFKHITPLOSKEVAIRIFOGCOFRSEVAEOETTEYAKN 300
QY 269 IPGFVNLNDQVTLTKYGVHEIITMLASLMNKDGVLSISGOGFMTREFLSLRKPFCD 328
Db 301 IPGFVNLNDQVTLTKYGVHEIITMLASLMNKDGVLSISGOGFMTREFLSLRKPFCD 360
QY 329 FMEKFEFAKFNALIEDDSDLAIFIAVITISGDRPGLLVKPIEDIQDNLQALEYOLK 388
Db 361 FMEKFEFAKFNALIEDDSDLAIFIAVITISGDRPGLLVKPIEDIQDNLQALEYOLK 420
QY 389 LNHESQSLFAKLQKMTDLROIVTEHVOLLQVITKKTETMSLHPLLOEYIKDLY 443
Db 421 LNHESQSLFAKLQKMTDLROIVTEHVOLLQVITKKTETMSLHPLLOEYIKDLY 475

```

```

RESULT 7
US-08-134-557D-2
; Sequence 2, Application US/08134557D
; Patent No. 6200802
; GENERAL INFORMATION:
; APPLICANT: Greene, Marianne E.
; TITLE OF INVENTION: Human Peroxisome Proliferator Activated
; TITLE OF INVENTION: Receptor Gamma: Compositions and Methods
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
; STREET: 2 Prudential Plaza, Suite 4700 180 N. Stetson
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/134, 557D
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Katz, Martin L.
; REGISTRATION NUMBER: 25,011
; REFERENCE/DOCKET NUMBER: ARCH:098
; TELEPHONE: 312-616-5400
; TELEFAX: 312-616-5460
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-134-557D-2

```

```

Query Match 58.3%; Score 2044.5; DB 4; Length 476;
Best Local Similarity 85.9%; Pred. No. 1.1e-169;
Matches 409; Conservative 2; Mismatches 2; Indels 63; Gaps 2;

```

```

QY 31 MVDEMPFMTNFGISSVDLSVMEHSHSPDIKPFVTVDSSISAPHYEDIPFTRADPMV 90
Db 1 MVDEMPFMTNFGISSVDLSVMEHSHSPDIKPFVTVDSSISAPHYEDIPFTRADPMV 60
QY 91 ADVKYDLKLOEYOSAIKVEBPASPPYSEKTOLYN-----124
Db 61 ADVKYDLKLOEYOSAIKVEBPASPPYSEKTOLYNRPHEEPSNSLMAIECRVCGDKASGFH 120
QY 125 -----RNKCQYCRFOKCLAVGSMHNAIRF 148
Db 121 YGVHACGCKGFFRRTIRLKLIVDRCDLNCRIHKSSRNKCQYCRFOKCLAVGSMHNAIRF 180

```

QY 149 GRMPOAEKELLAETSSDIDOLNPESADLR-ALAKHLVDSYIKSPFLTKAKARAILTGKT 207
11: |||||
Db 181 GRIAOAEKEKLLAETSSDIDOLNPESADLRQALAKHLVDSYIKSPFLTKAKARAILTGKT 240
QY 208 TDKSPFIYDMSLMMGEDKIKFKHITPLQEOSKEVAIRIFOGCQFRSVAVOETEVAK 267
|||
Db 241 TDKSPFIYDMSLMMGEDKIKFKHITPLQEOSKEVAIRIFOGCQFRSVAVOETEVAK 300
QY 268 STPGFVNLNDQVTLTKYGVHEIITMLASLNNKQVLISSGQGMTRFEFLSKRPFG 327
|||
Db 301 STPGFVNLNDQVTLTKYGVHEIITMLASLNNKQVLISSGQGMTRFEFLSKRPFG 360
QY 328 DMEPKFEFAVKFNALDELDSDLAIFIAVILISGDRPGILNVKPIDIDNLLQALELQ 387
|||
Db 361 DMEPKFEFAVKFNALDELDSDLAIFIAVILISGDRPGILNVKPIDIDNLLQALELQ 420
QY 388 LNHPESSQLFAKLQKMTDLROIVTEHVOLLQVIRKKTETDMSLHPLLQEIYKDL 443
|||
Db 421 LNHPESSQLFAKLQKMTDLROIVTEHVOLLQVIRKKTETDMSLHPLLQEIYKDL 476

RESULT 8
US-08-764-870-10
; Sequence 10, Application US/08764870
; Patent No. 6236946
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Thomas S
; APPLICANT: Baxter, John D
; APPLICANT: Fletterick, Robert J
; APPLICANT: Wagner, Richard L
; APPLICANT: Kushner, Peter J
; APPLICANT: Aprilletti, James W
; APPLICANT: West, Brian
; TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand
; TITLE OF INVENTION: Binding Domains
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward
; STREET: Five Palo Alto Square, 3000 El Camino Real
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/764,870
; FILING DATE: 13-DEC-1996
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/008,540
; FILING DATE: 13-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/008,543
; FILING DATE: 13-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/008,606
; FILING DATE: 14-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Nakamura, Jackie N
; REGISTRATION NUMBER: 35,966
; REFERENCE/DOCKET NUMBER: UCAL-246/01US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)843-5000
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 475 amino acids
; TYPE: amino acid
; STRANDEDNESS:

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-764-870-10
Query Match 58.2%; Score 2043; DB 4; Length 475;
Best Local Similarity 84.8%; Pred. No. 1.5e-169;
Matches 403; Conservative 6; Mismatches 4; Indels 62; Gaps 1;
QY 31 MVDTEPFWPTNFGTSSVDSLVMSHDSHSDIKPFTVDSSTPHYDIPFTTRDPVY 90
|||
Db 1 MVDTEPFWPTNFGTSSVDSLVMSHDSHSDIKPFTVDSSTPHYDIPFTTRDPVY 60
QY 91 ADYKYDLKLOEYOSAKIVPASPPIYSEKAQLYNNRHEEPSLSMAIECRVCGDKASGH 124
|||
Db 61 ADYKYDLKLOEYOSAKIVPASPPIYSEKAQLYNNRHEEPSLSMAIECRVCGDKASGH 120
QY 125 -----RNKQYCRPQKCLAVGSHNAIRF 148
|||
Db 121 YGVHACEGCKGFERRTIRLKLIVRCDLNCRIHKKSNNKQYCRPQKCLAVGSHNAIRF 180
QY 149 GRMPOAEKELLAETSSDIDOLNPESADLRALAKHLVDSYIKSPFLTKAKARAILTGKT 208
|||
Db 181 GRMPOAEKELLAETSSDIDOLNPESADLRALAKHLVDSYIKSPFLTKAKARAILTGKT 240
QY 209 DKSPFIYDMSLMMGEDKIKFKHITPLQEOSKEVAIRIFOGCQFRSVAVOETEVAKS 268
|||
Db 241 DKSPFIYDMSLMMGEDKIKFKHITPLQEOSKEVAIRIFOGCQFRSVAVOETEVAKN 300
QY 269 IGFVNLNDQVTLTKYGVHEIITMLASLNNKQVLISSGQGMTRFEFLSKRPFGD 328
|||
Db 301 IGFVNLNDQVTLTKYGVHEIITMLASLNNKQVLISSGQGMTRFEFLSKRPFGD 360
QY 329 FMEPKFEFAVKFNALDELDSDLAIFIAVILISGDRPGILNVKPIDIDNLLQALELQ 388
|||
Db 361 FMEPKFEFAVKFNALDELDSDLAIFIAVILISGDRPGILNVKPIDIDNLLQALELQ 420
QY 389 LNHPESSQLFAKLQKMTDLROIVTEHVOLLQVIRKKTETDMSLHPLLQEIYKDL 443
|||
Db 421 LNHPESSQLFAKLQKMTDLROIVTEHVOLLQVIRKKTETDMSLHPLLQEIYKDL 475

RESULT 9
US-08-980-115-10
; Sequence 10, Application US/08980115
; Patent No. 6266622
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Thomas S.
; APPLICANT: Baxter, John D.
; APPLICANT: Fletterick, Robert J.
; APPLICANT: Wagner, Richard L.
; APPLICANT: Kushner, Peter J.
; APPLICANT: Aprilletti, James W.
; APPLICANT: West, Brian L.
; APPLICANT: Shiau, Andrew K.
; TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS
; FILE REFERENCE: UCAL-246/02US
; CURRENT APPLICATION NUMBER: US/08/980,115
; CURRENT FILING DATE: 1997-11-26
; EARLIER APPLICATION NUMBER: 08/764,870
; EARLIER FILING DATE: 1996-12-13
; EARLIER APPLICATION NUMBER: 60/008,606
; EARLIER FILING DATE: 1995-12-14
; EARLIER APPLICATION NUMBER: 60/008,543
; EARLIER FILING DATE: 1995-12-13
; EARLIER APPLICATION NUMBER: 60/008,540
; EARLIER FILING DATE: 1995-12-13
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 10
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:

NAME/KEY: DOMAIN
LOCATION: (203)-(475)
OTHER INFORMATION: minimal ligand binding domain
US-08-980-115-10

Query Match 58.2%; Score 2043; DB 4; Length 475;
Best Local Similarity 84.8%; Pred. No. 1.5e-169;
Matches 403; Conservative 6; Mismatches 4; Indels 62; Gaps 1;

```

QY 31 MDTETMPPTNGISSVSLVSHSHSPDIKPFPTVDSSTPHHYDIPPTRPDPV 90
DB 1 MDTETMPPTNGISSVSLVSHSHSPDIKPFPTVDSSTPHHYDIPPTRPDPV 60
QY 91 ADYKYLKLOEYQSAIKVBPASPPYSEKTOLYN----- 124
DB 61 ADYKYLKLOEYQSAIKVBPASPPYSEKTOLYNRPHEPNSLMAIECRVCGDKASGFH 120
QY 125 -----RNKCYCRFOKCLAVGSHNAIRF 148
DB 121 YGVHAGECGKGFERRTRIKLIYDRCDLNCRIHKSRNKCQYCRFOKCLAVGSHNAIRF 180
QY 149 GRMPOAEKELLAETISSDIDOLNPESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTT 208
DB 181 GRMPOAEKELLAETISSDIDOLNPESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTT 240
QY 209 DKSPFVITDMNSLMGDEKIKFKHITPLOSKEVAIRIFOGCOFRSVAVOETETAKS 268
DB 241 DKSPFVITDMNSLMGDEKIKFKHITPLOSKEVAIRIFOGCOFRSVAVOETETAKN 300
QY 269 TGFVNLNDNDVTLTKYGVHEITITMLASLNNKDGVLISEGOGFMTREFLSLRKPPGD 328
DB 301 TGFVNLNDNDVTLTKYGVHEITITMLASLNNKDGVLISEGOGFMTREFLSLRKPPGD 360
QY 329 FMEPKFEFVKNALDLSDLAIFAVITILSGDRPGLNVKPIEDIONLLOALELQK 388
DB 361 FMEPKFEFVKNALDLSDLAIFAVITILSGDRPGLNVKPIEDIONLLOALELQK 420
QY 389 LNHPESSQFAKLQKMTDLROITVEHVOLLQYIKKTEETDMSLHPLLOEIKYKDL 443
DB 421 LNHPESSQFAKLQKMTDLROITVEHVOLLQYIKKTEETDMSLHPLLOEIKYKDL 475

```

RESULT 10
US-08-764-870-9
Sequence 9, Application US/08764870
Patent No. 6236946
GENERAL INFORMATION:
APPLICANT: Scanlan, Thomas S
APPLICANT: Baxter, John D
APPLICANT: Fletterick, Robert J
APPLICANT: Wagner, Richard L
APPLICANT: Kushner, Peter J
APPLICANT: Apriletti, James W
APPLICANT: West, Brian
TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand
TITLE OF INVENTION: Binding Domains
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooley Godward
STREET: Five Palo Alto Square, 3000 El Camino Real
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/764,870
APPLICATION NUMBER: US/08-764,870
FILING DATE: 13-DEC-1996
CLASSIFICATION: 530

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,540
FILING DATE: 13-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,543
FILING DATE: 13-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/008,606
FILING DATE: 14-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Nakamura, Jackie N
REGISTRATION NUMBER: 35,966
REFERENCE/DOCKET NUMBER: UCAL-246/0105
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)843-5000
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 441 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-764-870-9

Query Match 31.4%; Score 1101.5; DB 4; Length 441;
Best Local Similarity 63.8%; Pred. No. 9.8e-88;
Matches 208; Conservative 56; Mismatches 61; Indels 1; Gaps 1;

```

QY 119 KQOLYRNKCQYCRFOKCLAVGSHNAIRFGMRPOAEKELLAET-SSDIDOLNPESADL 177
DB 116 KQKNNRNKCQYCRFOKCLAVGSHNAIRFGMRPEAEKRLVAGLVANGSOYNQVADL 175
QY 178 RALAKHLYDSYIKSFPLTKAKARAILTGKTTDKSPFVITDMNSLMGDEKIKFKHITPQ 237
DB 176 KAFSKHITVAYLKNFMTKKARSITGKASHAPRVHDIETLMOAEGVLWKOLVNL 235
QY 238 EOSKEVAIRIFOGCOFRSVAVOETETAKSIPGFVNLNDNDVTLTKYGVHEITITMLA 297
DB 236 PPKKEISVHVFRCQTTVEYRELTEFAKSIPIFSFSLNDQVTLTKYGVHAIIFAMLA 295
QY 298 SLNNKDGVLISEGOGFMTREFLSLRKPPGDMEKFEFVKNALDLSDLAIFIAI 357
DB 296 SIYNNKDGVLVANGSGFVTEFELSLRKPSPDIIEPKFEFVKNALDLSDLAIFIAI 355
QY 358 IISGDRPGLNVKPIEDIONLLOALELQKLNHPESSQFAKLQKMTDLROITVEHVQ 417
DB 356 IICGDRPGLNVKPIEDIONLLOALELQKLNHPESSQFAKLQKMTDLROITVEHVQ 415
QY 418 LLOVYIKKTEETDMSLHPLLOEIKYKDL 443
DB 416 MMQRIKTEETDMSLHPLLOEIKYKDMY 441

```

RESULT 11
US-08-980-115-9
Sequence 9, Application US/08980115
Patent No. 6266622
GENERAL INFORMATION:
APPLICANT: Scanlan, Thomas S
APPLICANT: Baxter, John D
APPLICANT: Fletterick, Robert J
APPLICANT: Wagner, Richard L
APPLICANT: Kushner, Peter J
APPLICANT: Apriletti, James W
APPLICANT: West, Brian L
APPLICANT: Shiao, Andrew K
TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS
FILE REFERENCE: UCAL-246/0205
CURRENT APPLICATION NUMBER: US/08/980,115
CURRENT FILING DATE: 1997-11-26
EARLIER APPLICATION NUMBER: 08/764,870
EARLIER FILING DATE: 1996-12-13
EARLIER APPLICATION NUMBER: 60/008,606


```
STREET: 444 SOUTH FLOWER STREET, SUITE 2000
CITY: LOS ANGELES
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,200
FILING DATE: 07-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: REITER, STEPHEN E.
REGISTRATION NUMBER: 31,192
REFERENCE/DOCKET NUMBER: P41 9971
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-546-1995
TELEFAX: 619-546-9392
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 440 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-200-4
```

```
Query Match 30.7%; Score 1076.5; DB 2; Length 440;
Best Local Similarity 62.3%; Pred. No. 1.5e-85;
Matches 203; Conservative 58; Mismatches 64; Indels 1; Gaps 1;
```

```
119 KTLVLRNKKCYCRFOKCLAVGSHNAIRGRMPQAEKELAEI-SSDIDOLNPSADL 177
115 KIKKRNKRCQYRFOKCLALGSHNAIRGRMPDGEKRLVAGLASECCQNPQADL 174
178 RALAKHLYDSYKISFPLTKAKARAILTGKTTKSPFVIYDMSLMGEGDKIKFKHITPLQ 237
175 KAKSKHLYNAYLKNFMNTKKARSILTKGSSHNAFPVIHDIETLMQAEKGLVWKQVNLG 234
238 EDSKEVAIRFOGQCFRSVAEAOEITEYAKSIGFVNLNDNDVYTLTKGVHIIITMYLA 237
235 PRYNEISVHFYRCQSTVRELTETFAKNIPNFSSFLFNDQVTLTKGVHNAIFAMLA 234
298 SLANKGVLISGQGFMTREFTLSLRKPPGDFMEPKFEFAVKFNALDELDDSLAIFAVI 357
295 SLVKNKGGLVANSQVTHFELSLKRPSPDIIEPKFEFAVKFNALDELDDSLAIFAAI 354
358 IISGDRPGLLVKPIEDIDNLQALELQELKLNHPRESSQLFAKLQKMTDLQIOTEHVAQ 417
355 ILGDRPGLLVNVPQVBAIQITILRALEPHLVNHPDSQYLFRLQKMDLRQLVTEHNAQ 414
418 ILQVYIKTEIDMSLHPLQEIYNDLY 443
415 MMQMLKTESETLHPLQEIYKDMY 440
```

```
RESULT 14
US-08-459-287-2
Sequence 2, Application US/08459287
Patent No. 5686596
GENERAL INFORMATION:
APPLICANT: Mukherjee, Ranjan
TITLE OF INVENTION: HUMAN PEROXISOME PROLIFERATOR ACTIVATED
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 611 West Sixth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
```

```
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM compatible
OPERATING SYSTEM: IBM P.C. DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/459,287
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/143,215
FILING DATE: October 25, 1993
APPLICATION NUMBER: 08/141,500
FILING DATE: October 22, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 204/138
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 468 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-459-287-2
```

```
Query Match 30.7%; Score 1076.5; DB 1; Length 468;
Best Local Similarity 49.8%; Pred. No. 1.6e-85;
Matches 226; Conservative 65; Mismatches 106; Indels 67; Gaps 9;
```

```
31 MVDTEPMPFTN-----FGISSVDLSVMDHSHSPDIRFTVY---- 68
1 MVDTEPPLCPLEAGDLESPLSEFLOEMGNIQISIGDSSGSGFTEYQYLSC 60
69 ---DFSSIS-TPHYEDIPFTRDPVADYKYDLQIYES-AIKVE-----PASPPY- 116
61 PGSDGSVITDLSPASPSSTVTPVPG-----SVDESPSGALNIECRICGDKASGYHG 115
117 -----SEKTOLYNNKKCYCRFOKCLAVGSHNAIRFGR 150
116 VHAECKGFEFRRTIRLKYDCKDSCKIQKKRNKKCYCRFHKCLSGMSHNAIRFGR 175
151 MPQAEKELAEI-SSDIDOLNPSADLRALAKHLYDSYKISFPLTKAKARAILGKTTD 209
176 MPRESAKLKAETLTCEHDI EDSERVADLSLAKRIYEAFLKNMKNKVARAVYILSGASN 235
210 KSPFVIYDMSLMGEGDKIKFKHITPLQESKEVAIRIFOGQCFRSVAEAOEITEYAKSI 269
236 NRPFVYHDMETLCKMAKTLVAK-LVANGIONKEAEVRIFHCOCQTSVEVYTELTERAKAI 294
270 PGFVNLDNDQVTLTKGVHIIITMYLASLMKDGVLISGQGFMTREFLSLRKPPGDF 329
295 PGFANDLNDQVTLTKGVYEAIFAMLSVMNKGDLVAVYNGCFIRREFLSLRKPPCDI 354
330 MEKFEFAVKFNALDELDDSLAIFAVIILSGRPGGLLVKPIEDIDNLQALELQELK 389
355 MEKFEFAVKFNALDELDDSLISLEVAIICCGDRPGLLVNGHIEKQESIVHNLRLHDS 414
390 NHPRESSQLFAKLQKMTDLQIOTEHVAQVYIKTEIDMSLHPLQEIYNDLY 443
415 NHDDTLFEPKLLQKMDLRQLVTEHNAQLVQIITKIESDALHPLQEIYKDMY 468
```

```
RESULT 15
US-08-764-870-8
Sequence 8, Application US/08764870
Patent No. 6236946
GENERAL INFORMATION:
APPLICANT: Scanlan, Thomas S
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 25, 2003, 04:00:36 ; Search time 69 Seconds
(without alignments)
3057.878 Million cell updates/sec

Title: US-09-931-007A-1

Perfect score: 3508
Sequence: 1 MGFTLGDSPIDPESDSFTDT.....KTETDMSLHPLQLQETRYDLY 688

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO/US09931007/runat_14022003_102554_4697/app_query.fasta_1.839
-DB=Issued_Patents.NA -QEXT=fastac -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MILK=0 -MAXLEN=2000000000
-USER=US09931007.ecgn.1.1.32.etrnat.14022003_102554_4697 -NCP=6 -ICPU=3
-NO_XIPYX -NO_MMAP -LARGEDQUERY -NEG_SCORES=0 -WAIT -LONGLOGS -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCUTS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2231	63.6	1508	4	US-09-128-142-3
2	2231	63.6	1618	3	Sequence 3, Appli
3	2196.5	62.6	1679	4	US-09-484-345-3
4	2089	59.5	1434	4	US-09-514-247A-5
5	2065.5	58.9	1844	4	US-09-128-142-1
6	2065	58.9	1796	4	US-08-134-357D-1
7	2065	58.9	1796	4	US-09-255-392-1
8	2065	58.9	2005	2	US-08-484-200-1
9	2065	58.9	2005	3	US-08-477-493-1
10	2065	58.9	2005	3	US-08-465-375-1
11	1076.5	30.7	2012	4	US-09-788-070-1
12	1071	30.5	2009	2	US-08-484-200-3
					Sequence 9, Appli

13	1071	30.5	2009	1	US-08-463-694-9	Sequence 9, Appli
14	1071	30.5	2009	1	US-08-694-501-9	Sequence 9, Appli
15	1057.5	30.1	1407	1	US-08-459-287-1	Sequence 1, Appli
16	521	14.9	373	3	US-08-917-653-3	Sequence 3, Appli
17	341.5	9.7	1952	1	US-08-333-358-1	Sequence 3, Appli
18	341.5	9.7	1952	1	US-08-463-694-1	Sequence 1, Appli
19	341.5	9.7	1952	1	US-08-694-501-1	Sequence 1, Appli
20	333	9.5	277	3	US-08-917-653-4	Sequence 4, Appli
21	326.5	9.3	2989	6	5232606-1	Patent No. 5232606
22	321.5	9.2	3511	3	US-08-893-747-13	Sequence 13, Appli
23	315	9.0	2928	2	US-08-095-728B-3	Sequence 3, Appli
24	315	9.0	2928	5	PCT-US92-02320A-3	Sequence 3, Appli
25	315	9.0	2940	2	US-08-592-383-1	Sequence 1, Appli
26	315	9.0	2940	6	5171671-1	Patent No. 5171671
27	315	9.0	3036	1	US-08-306-691B-52	Sequence 52, Appli
28	315	9.0	3036	5	US-08-095-728B-1	Sequence 1, Appli
29	315	9.0	3036	2	PCT-US92-02320A-1	Sequence 1, Appli
30	313	8.9	1576	6	5260432-1	Patent No. 5260432
31	313	8.9	1659	1	US-08-333-358-7	Sequence 7, Appli
32	305	8.7	1659	1	US-08-463-694-7	Sequence 7, Appli
33	305	8.7	1659	1	US-08-694-501-7	Sequence 7, Appli
34	293.5	8.4	2658	2	US-08-592-383-3	Sequence 7, Appli
35	289.5	8.3	2043	1	US-07-737-736B-6	Sequence 6, Appli
36	283	8.1	1649	2	US-08-466-120-1	Sequence 1, Appli
37	283	8.1	1649	5	PCT-US94-07266-1	Sequence 1, Appli
38	281.5	8.0	1677	2	US-08-372-652-14	Sequence 14, Appli
39	281.5	8.0	1677	5	PCT-US95-16311-14	Sequence 14, Appli
40	278.5	7.9	1959	1	US-08-342-411A-3	Sequence 3, Appli
41	278.5	7.9	1959	5	PCT-US94-12883-4	Sequence 4, Appli
42	278	7.9	2231	1	US-08-496-631-1	Sequence 1, Appli
43	276.5	7.9	1662	1	US-08-336-408B-5	Sequence 5, Appli
44	276.5	7.9	1662	5	PCT-US91-00399-5	Sequence 5, Appli
45	276.5	7.9	1860	2	US-08-372-652-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-128-142-3
Sequence 3, Application US/09128142
Patent No. 6294559
GENERAL INFORMATION:
APPLICANT: Smith, Roy G.
TITLE OF INVENTION: ANTIPROLIFERATIVE AGENTS ASSOCIATED WITH
PEROXISOME PROLIFERATOR ACTIVATED RECEPTORS GAMMA AND G
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Jack L. Tribble
STREET: 126 E. Lincoln Ave., P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: Power Macintosh 7500/100
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/128,142
FILING DATE: 03-Aug-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/844,007
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Tribble, Jack L.
REGISTRATION NUMBER: 32,633
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-5321
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 3:

```

;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1518 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-128-142-3

Alignment Scores:
Pred. No.:      6,986-252      Length:      1518
Score:          2231.00      Matches:      443
Percent Similarity: 87.72%      Conservative: 0
Best Local Similarity: 87.72%      Mismatches: 0
Query Match:      63.60%      Indels:      62
DB:               4           Gaps:      1

US-09-931-007a-1 (1-688) x US-09-128-142-3 (1-1518)

QY 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThr 20
   |||||
DB 1 ATGGGTGAACACTGGAGCATTCCTCTATTGACCCGAAAGCATTCCTTCACTGATACA 60
QY 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheTrpPro 40
   |||||
DB 61 CTGTCTGCAAAACATATACACAAAGAACCATGCTTGACACAGAGATGCCATTTCTGGCCC 120
QY 41 ThrAsnPhgGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
   |||||
DB 121 ACCAACTTTGGATGACCTCCGCTGATCTCTCGTATGGAAGACCACTCCACTCTCTTT 180
QY 61 AspIleGlyProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80
   |||||
DB 181 GATATCAACCCCTTCACTGCTTGTGACTTCTCCAGCATTTTACTCCACATTAACGAAGAC 240
QY 81 IleProPheThrArgThrAspProValValAlaAspTyrIleTyrAspLeuGln 100
   |||||
DB 241 ATTTCATTTACAAAGAACATCCAGTGTGAGATTACAAAGTATGACCTGAAACTTCAA 300
QY 101 GlnTyrGlnSerAlaIleGlyValGluProAlaSerProTyrTyrSerGluTyrThr 120
   |||||
DB 301 GAGTACCAAGAGCAATCAATCAAGAGCTGCATCTCCACTTATATTCTGAGAAAGACT 360
QY 121 GlnLeuTyrAsn----- 124
   |||||
DB 361 CAGCTCTCAATAGCCATGATGAAGACCTTCCAACTCCCTCATGSCAATGAAATGCTGT 420
QY 124 ----- 124
DB 421 GTCTGTGAGATAAAGCTTCTGATTTCACTATGAGATTCAATGCTTGTGAAGATGCAAG 480
QY 124 ----- 124
DB 481 GGTTCCTCCGAGAACATCAATGGAAGCTTATCTATGACAGATGATGATCTTAACTGT 540
QY 125 ----- 125
   |||||
DB 541 CGGATCCACAAAAAGATGAATTAATGCTAGTACTGCTCGGTTTCAGAAATGCCCTGCA 600
QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluGlnGluTyr 158
   |||||
DB 601 GTGGGATGTCATTAATGCCATCAGGTTTGGGCGGATGCCACAGGCCGAGAGGAAG 660
QY 159 LeuLeuAlaGlnIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
   |||||
DB 661 CTGTTGGCGGAATCTCCAGTATATGACACGAGTGAATCCAGATGCCCTACCTCCGG 720
QY 179 AlaLeuAlaValHisLeuTyrAspSerTyrIleGlySerPheProLeuThrLysAlaLys 198
   |||||
DB 721 GCCCTGGCAAAACATTTGATGACATCATATAAAGTCTTCCCGGTGACCAAGGAAG 780
QY 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218
   |||||
```

```

DB 781 GCGAGGCGCATCTTGCACAGAAAGACAAACAGACAAATGCATTCCTTATCTATGACATG 840
QY 219 AsnSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGlu 238
   |||||
DB 841 AATTCTTATATATGAGGAGAAATAAATCAAGTTCAAAACATCATCAACCCCTCGACAGAG 900
QY 239 GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258
   |||||
DB 901 CAGAGCAAGAGAGTGCGCCATCCGATCTTTCAGGGCTGGCAGATTTCGCTCCGTGAGAGCT 960
QY 259 ValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278
   |||||
DB 961 GTCCAGGAGATACACAGAGATTCCTGCTGTTTGTAAATCTTGACTTTGAAAC 1020
QY 279 AspGlnValThrLeuLeuTyrGlyValHisGluIleIleTyrThrMetLeuLysSer 298
   |||||
DB 1021 GACCAAGTATCTCTCAATATGAGTCCAGAGATCATTTATACACATGCTGACCTCC 1080
QY 299 LeuMetAsnLysAspGlyValLeuIleSerGlnGlnGlyPheMetThrArgGluPhe 318
   |||||
DB 1081 TTGATGAATAAAGATGGGCTTCTCATATCCGAGGCCAAAGCTTCATGACAAAGGAGTTT 1140
QY 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338
   |||||
DB 1141 CTAAAGAGCCTCCGAAAGCCTTTTGTGACTTTATGAGGCCCAAGTTTGAGTTTCTGTG 1200
QY 339 LysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaIleIle 358
   |||||
DB 1201 AAGTTCATGCACTGGAATTAATGACAGCAGCTTGGCAATTTTATTGCTCTCATATTAT 1260
QY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
   |||||
DB 1261 CTCAGAGGAGACCGCCGAGGTTTGTGATGTGAAGGCCATTTGAAGACATTCACAAACAAC 1320
QY 379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerLysLeuPhe 398
   |||||
DB 1321 CTGCTTCAAGCCCTGAGCTCCAGCTGAAAGTGAACACCCCTGAGCTCCACAGCTGT 1380
QY 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeu 418
   |||||
DB 1381 GCCAGCTGCTCCAGAAATGACAGACCTCAGACAGATTTGTACGGAACAGCTGCAGCTA 1440
QY 419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIle 438
   |||||
DB 1441 CTGCAAGTGATCAAGAACGAGACAGACATGAGTCTTCAACCCCTCCGACGAGATC 1500
QY 439 TyrLysAspLeuTyr 443
   |||||
DB 1501 TACAAGGACTGTATC 1515

RESULT 2
US-09-484-345-3
; Sequence 3, Application US/09484345
; Patent No. 6159734
; GENERAL INFORMATION:
; APPLICANT: Robert McKay
; APPLICANT: Alexander H. Borchers
; APPLICANT: Brenda F. Baker
; TITLE OF INVENTION: ANTISENSE MODULATION OF PEROXISOME PROLIFERATOR-ACTIVATED RECE
; FILE REFERENCE: R1S-0104
; CURRENT APPLICATION NUMBER: US/09/484, 345
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO: 3
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)...(1608)
US-09-484-345-3

Alignment Scores:
```


DB: 4 Gaps: 2
US-09-931-007A-1 (1-688) x US-09-514-247A-5 (1-1679)
QY 1 MetGlyGlnThrLeuGlnLysSerProIleAspProGlnSerAspSerPheThrAspThr 20
DB 159 ATGGGGAAACTCTGGGAATTCCTATTGACCCAGAAAGCATTCCTCACTGATACAA 218
QY 21 LeuSerAlaAsnIleSerGlnLumethrMetValAspThrGlnumethrProthleppro 40
DB 219 CTGTCTGCAACATATACAAAGAAATGACCATGTTGACACAGAGATGCGCATTTCTGCCC 278
QY 41 ThrAspPheGlyIleSerSerValAspLeuSerValMetGlnAspHisSerHisSerPhe 60
DB 279 ACCAACTTGGGATCAGCTCCGGATCTCTCCGTAATGGAAGACCACTCCACCTCTTT 338
QY 61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisThrGluAsp 80
DB 339 GATATCAAGCCCTTCACCTACTGTTGACTTCCAGCATTTTCACATTCACATTAACGAAGAC 398
QY 81 IleProPheThrArgThrAspProValValAlaAspTyrIleTyrAspLeuLysLeuGln 100
DB 399 ATTCCTTACCAAGAACATCCAGTGGTTCAGATTTACAAAGATGACCTGAACACTTCAA 458
QY 101 GlnTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThr 120
DB 459 GAGTACCAAGGTGCAATCAAAAGTGAGCCTGCATCTCCACCTTATATTCTGGAAGACT 518
QY 121 GlnLeuTyrAsn----- 124
DB 519 CAGCTCTACAAATTAAGCCTCATGAGAGCCTTCCAACTCCCTCATGGCAATTGAATGTCGT 578
QY 124 ----- 124
DB 579 GTCGTGAGATAAAGCTCTGGATTTCACTATGAGAGTTCATGCTTGTAAGATGCAG 638
QY 124 ----- 124
DB 639 GGTTTCTTCGGAACAATCAATGAACTTATCTATGACAGATGATCTTAACGT 698
QY 125 ----- -ArgAsnLysCysGlnIntyrCysArgPheGlnLysCysLeuAla 138
DB 699 CGGATCCACAAAAAGTGAATTAATGTCAGTACTGTCGGTTTCAGAAAGCCCTGGA 758
QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaLysGluLys 158
DB 759 GTGGGGATGTCATATATGCCATCAAGTTTGGGGGATGCGACAGGCCAGAGAGAGAG 818
QY 159 LeuLeuAlaGlnIleSerSerAspIleAspGlnLeuAsnProGlnSerAlaAspLeuArg 178
DB 819 CTGTTGGCGGAGATCTCCAGTGATATCGACGATCAATCCAGAGTCCGCTGACCTCCGT 878
QY 179 ---AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAla 197
DB 879 CAGGCCCTGGCAAAACATTGTATGACTCATATCAATAAATCCCTCCCGCTGACCAAGGA 938
QY 198 LysAlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAsp 217
DB 939 AAGGGAGGGGATCTGTGACAGGAAAGACACAGACAAATCACATTCCTTATCATATGAC 998
QY 218 MetAsnSerLeuMetMetGlyLysAspLysIleLysPheLysHisIleThrProLeuGln 237
DB 999 ATGAATTCCTTATGATGAGGAGAAATAAATCAAGTTCAAAACATCACCCCTCGCAG 1058
QY 238 GlnGlnSerLysGlnValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGlu 257
DB 1059 GAGCAGAGCAAGAGGTGGCCATCCGCACTTTCAGGGCTGCACAGTTTCGCTCCGTGGAG 1118
QY 258 AlaValGlnGlnIleThrGlnTyrAlaLysSerIleProGlyPheValAsnLeuAspLeu 277
DB 1119 GCTGTGCAAGAAATCATCAGAGATATGCCAAAAGCATTCCTGGGTTTGTAAATCTTGACTTG 1178
QY 278 AsnAspGlnValThrLeuLeuLysTyrGlyValHisGlnIleIleTyrThrMetLeuAla 297

DB 1179 AAGACACAGTAACCTCTCTCAAAATATGAGATCCAGACATGATTTACACATGCTGGCC 1238
QY 298 SerLeuMetAsnLysAspLysValLeuIleSerGlnGlyGlnGlyPheMetThrArgGln 317
DB 1239 TCCTTGATGAATAAAGATGGGTCTCATATATCCGAGGGCAAGGCTTCAATGACAAAGGAG 1298
QY 318 PheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAla 337
DB 1299 TTTCATAAGACCTCGGAAGCCTTTTGGTGACTTATATGAGACCAAGTTTACGTTTGGT 1358
QY 338 ValLysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIle 357
DB 1359 GTGAAGTCAATGACATCGAATTAAGATGACAGACGACTTGGCAATGATTTATCTGTCATTT 1418
QY 358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGlnAspIleGlnAsp 377
DB 1419 ATTCATAGTGAAGACCCGACAGTTTGTGAAATGTAACCCCATTTGAACATTCACAGAC 1478
QY 378 AsnLeuGlnAlaLeuGlnLeuGlnLeuLysIleAsnHisProGlnSerSerGlnLeu 397
DB 1479 AACCTGCTCAAGCCCTGAGCTCCAGCTGAAGCTGAACCATCCTGAGTCTCACAGCTG 1538
QY 398 PheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGln 417
DB 1539 TTTGCCAAGCTGCTCCAGAAATGACAGACCTCAGACAGATTTGTACGGAACACGCTGCAG 1598
QY 418 LeuLeuGlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeuGlnGln 437
DB 1599 CTACTGCAGGTATCAAGAAGCGAGACAGACATGAGTTCACCCGCTCTCTGCAGAG 1658
QY 438 IleTyrLysAspLeuTyr 443
DB 1659 ATCTACAAGACTTGATC 1676
RESULT 4
US-09-128-142-1
; Sequence 1, Application US/09128142
; Patent No. 6294559
; GENERAL INFORMATION:
; APPLICANT: Smith, Roy G.
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENTS ASSOCIATED WITH
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Jack L. Tribble
; STREET: 126 E. Lincoln Ave., P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Power Macintosh 7500/100
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,142
; FILING DATE: 03-Aug-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,007
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Tribble, Jack L.
; REGISTRATION NUMBER: 32,633
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-5321
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1434 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-128-142-1

Alignment Scores:
Pred. NO.: 2.88e-235 Length: 1434
Score: 2089.00 Matches: 415
Percent Similarity: 87.00% Conservative: 0
Best Local Similarity: 87.00% Mismatches: 0
Query Match: 59.55% Indels: 62
DB: 4 Gaps: 1

US-09-931-007a-1 (1-688) x US-09-128-142-1 (1-1434)

QY 29 MetThrMetValAspThrGluMetProPheThrProThrAsnPheGlyIleSerSerVal 48
Db 1 ATGACCATGTTGACACAGAGATGCCATTCTGGCCACCACTTGGGATCAGCTCCGTG 60
QY 49 AspleuSerValMetGluAspHisSerHisSerPheAspIleIysProPheThrVal 68
Db 61 GATCTCTCCGAAATGGAGACCACTCCACCTCTTGATATCAAGCCCTCAGTACTGTT 120
QY 69 AsPheSerSerIleSerThrProHisThrGluAspIleProPheThrArgThrAspPro 88
Db 121 GACTTCTCCACCATTTCTACTCCACATTACGAGACATTCATTCCACAGAACAGATCCA 180
QY 89 ValValAlaAspTyrIlystYrAspleuLysLeuGlnIlyrGlnSerAlaIleIysVal 108
Db 181 GTGGTTGCAGATTACAACTATGACCTGAAGTCAAGATCAAGTCAATCAAGT 240
QY 109 GluProAlaSerProProTyrTyrSerIulysThrGlnLeuTyrAsn----- 124
Db 241 GAGCTGCATCTCCACCTTATTATCTGAGAACTCAGCTCTCAATAAGCCCTCATGAA 300
QY 124 ----- 124
Db 301 GAGCCTTCCAACTCCCTCATGATGCAATTGATGTCTGTCTGGAGATTAAGCTTTTGA 360
QY 124 ----- 124
Db 361 TTTCACATGAGATTGATGTTGTAAGATGCAAGGTTTCTTCCGAGAAATCAGA 420
QY 125 -----ArgAsn 126
Db 421 TTGAAGCTTATCTATGACAGATGTGATCTTAACCTGCGATCCACAAAAAAGTGAAT 480
QY 127 LysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAlaIle 146
Db 481 AAATTCAGTACTGTGGTTTTCAGAAATGCTTGCAGTGGGAGATGTCATTAATGACCATC 540
QY 147 ArgPheGlyArgMetProGlnAlaGluLysGluLysLeuAlaGluIleSerSerAsp 166
Db 541 AGGTTTGGCGGATGCCACAGCGGAGAAAGAGTGTGGCGGAGATCTCCAGTAT 600
QY 167 IleAspGlnLeuAsnProGlnSerAlaAspleuArgAlaLeuAlaIlyHisLeuTyrAsp 186
Db 601 ATCGACCGAGTGAATCGAGATCGCTGACCTCCGGGCCCTGGCAAAACATTGTATGAC 660
QY 187 SerTyrIleIysSerPheProLeuThrLysAlaIlyAlaValAlaIleLeuThrGlyLys 206
Db 661 TCATTCATTAAGTCTCTCCGCTGACCAACAAAGGAGGCGGATCTTACAGGGAAG 720
QY 207 ThrThrAspLysSerProPheValIleTyrAspMetAsnSerLeuMetGlyGluAsp 226
Db 721 ACAACAGACAAATCACCATTGCTATATGACATGAATTCCTTAATATGAGGAAGAT 780
QY 227 LysIleIysPheIlyHisIleThrProLeuGlnIlyGlnSerLysGluValAlaIleArg 246
Db 781 AAAATCAAGTTCAACACATCACCCTCGAGGAGGAGCAAGAAAGAGTGCCATCCGC 840

QY 247 IlePheGlnIlyCysGlnPheArgSerValGlnAlaValGlnIlyIleThrGluTyrAla 266
Db 841 ATCTTTACGGCTGCACAGTTCCTCCCTGGAGGCTGTGAGAGATCACAGATGATGCC 900
QY 267 LysSerIleProGlyPheValAsnLeuAspleuAsnAspGlnValThrLeuLysTyr 286
Db 901 AAAAGCATCTCTGTTTGTAAATCTTGACTTGAACGACCAAGTAACCTCTCAAAATAT 960
QY 287 GlyValHisGluIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyValLeu 306
Db 961 GGAGTCCACGATCATTTTACAAATGCTGCCCTCTTGATGAATTAAGATGGGCTTTC 1020
QY 307 IleSerGlnIlyGlnIlyPheMetThrArgGluPheLeuLysSerLeuArgLysProPhe 326
Db 1021 ATATCCGAGGCCCAAGGCTTATGACAAAGGAGTTCCTAAAGAGCTCCGAAAGCTTTT 1080
QY 327 GlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeuAsp 346
Db 1081 GGTGACTTATGAGAGCCCAAGTTGAGTTGCTGTGAAGTTCATGACACTGGAATTAAGAT 1140
QY 347 AsPheSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGlyLeu 366
Db 1141 GACAGCGACTTGGCAATATTATGCTGTCTATATCTCAGTGGAGAGACCGCCAGGTTTG 1200
QY 367 LeuAsnValLysProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGluLeuGln 386
Db 1201 CTGATATGGAAAGCCCATTTGAAGACATTCACAAACCTGCTACCAAGCCCTGAGCTCCAG 1260
QY 387 LeuLysLeuAsnHisProGluSerSerGlnLeuPheAlaLysLeuGlnLysMetThr 406
Db 1261 CTGAAGCTGAAACACACCTGAGTCTCCACAGCTGTTGGCAAGCTGCTCCAGAAATGACA 1320
QY 407 AspleuArgGlnIleValThrGluHisValGlnLeuGlnValIleLysLysThrGlu 426
Db 1321 GACCTCAGACAGATTGTCCGGAACACGTCGAGTACTGAGGATGATCAAGAACGAG 1380
QY 427 ThrAspMetSerLeuHisProLeuGlnIlyIleTyrLysAspLeuTyr 443
Db 1381 ACAGACATGAGTCTTACCCGCTCTCGAGAGATCTCAAGGACTGTATC 1431

RESULT 5
US-08-134-557D-1
Sequence 1, Application US/08134557D
Patent No. 6200802
GENERAL INFORMATION:
APPLICANT: Greene, Marianne E.
TITLE OF INVENTION: Human Peroxisome Proliferator Activated
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Rockey, Milanow & Katz, Ltd.
STREET: 2 Prudential Plaza, Suite 4700 180 N. Stetson
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/134,557D
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Katz, Martin L.
REGISTRATION NUMBER: 25,011
REFERENCE/DOCKET NUMBER: ARCH:098
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-616-5400
TELEFAX: 312-616-5460

```

; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1844 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 179..1606
;   US-08-134-557D-1
;
Alignment Scores:
Pred. No.: 2,46e-232      Length: 1844
Score: 2065.50           Matches: 414
Percent Similarity: 85.71%  Conservative: 6
Best Local Similarity: 84.49%  Mismatches: 7
Query Match: 58.88%       Indels: 63
DB: 4                     Gaps: 2
US-09-931-007A-1 (1-688) x US-08-134-557D-1 (1-1844)
Oy 23 AlaasnlieserGlnGluMethrMetValAspThrGluMetProPheThrProhTrsn 42
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 155 GCCGCCGTGGCCGCGAATATGCAATGGTGTGACACAGAGATCGATCTGGCCACCAC 214
Oy 43 PheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPheaspIle 62
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 215 TTTGGATCAGAGTCCCGTGGATCTCTCCGTAATGGAAAGCACTCCACCTCTTATATTC 274
Oy 63 LysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAspIlePro 82
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 275 AAGCCCTTCACTACAGTGGTACTCTCCAGCAATTTCTACTCCACATTTACAGAGACATTTCA 334
Oy 83 PheThrArgThrAspProValValAlaAspTyrIleTyrAspLeuLysLeuGlnGluTyr 102
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 335 TTCACAAACAAAGATCCAGTGTGTCACATTTCAAGTATGACCTGAACTTCAAGAGTAC 394
Oy 103 GlnSerAlaIleLysValGluProAlaSerProTyrTyrSerGluLysThrGlnLeu 122
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 395 CAAAGTCGAATCAAGTGGAGCTGTCATCTCCACCTATTTATTTGAGAGAGCTCAGCTC 454
Oy 123 TyrAsn----- 124
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 455 TACATTAAGCTCTCATGAAAGCCTTCCAACTCCCTCATGCAATTGAATGTCGTCTGT 514
Oy 124 ----- 124
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 515 GGAGATTAAGCTCTGATTTCTACTATGAGATTCATGCTTGTGAAGATGCAGAGGTTTC 574
Oy 124 ----- 124
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 575 TTCCGAGAACATTCAGATGAGTTTATGACAGATGATGATCTTAACGTGCGATC 634
Oy 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGly 140
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 635 CACAAAAAAGTAAATAAATGTCAGTACTGTGGTTTCAGAAATGCTTCGACGTGGGG 694
Oy 141 MetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGlnLysLeuLeu 160
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 695 ATGCTCTCATTAATGCCATCAGGTTGGGGGATGCCACAGGCCGGAAGGAAGAAACCTGTTG 754
Oy 161 AlaGluIleSerSerAspIleAspGlnLeuAsnProGlnSerAlaAspLeuArg---Ala 179
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 755 GCGAGATCTCTCAGATGATGACAGCAGTGAATCCAGAGTCCGCTGACCTCCGTCAGGCC 814
Oy 180 LeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLysAla 199
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 815 CTGCAAAACATTTGTTGACTCATACATAAAGTCTCCCGCTGACCAAAAGCAAGCGC 874
Oy 200 ArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMetAsn 219
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 875 AGGCGCATCTTGACAGGAAGACAAACAGACAAATCACCATTCTGTTATCTATGACATGAAT 934
```

```

Oy 220 SerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGln 239
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 935 TCCTTAATGATGGAGGAAGATTAATCAAGTTCAACACATCACCCCCCTGCAGAGCAG 994
Oy 240 SerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAlaVal 259
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 995 AGCAAAAGAGGTGGCCATCCGATCTTTCAGGGCTCCAGTTTCGCTCCGTGAGAGCTGTG 1054
Oy 260 GlnGluIleThrGlnTyrAlaLysSerIleProGlyPheValAsnLeuAsnAsnsp 279
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1055 CAGGAGATCACAGATATGCCCCAAAAGCAATTCCTGGTTTGTAAATCTTGATGAACGAC 1114
Oy 280 GlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeu 299
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1115 CAAGTAACTCTCCCTCAATATATGAGATCCACAGATCATTACAGATCTGCTGCTGTG 1174
Oy 300 MetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGluPheLeu 319
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1175 ATGAATTAAGATGGGGTCTCATATCCGAGGGCCCAAGGCTTCATGACAAAGGAGTTTCTA 1234
Oy 320 LysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaValLys 339
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1235 AAGAGCTGCGAAGACCTTTGTGTACTTTATGAGAGCCCAAGTTGATTTGCTGTGAAG 1294
Oy 340 PheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIleIleLeu 359
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1295 TTCATATGCATCGAATTTATGATGACAGCGACTTGGCAATATTTATTCGTCTATATTC 1354
Oy 360 SerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeu 379
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1355 AGTGAGAGACCCCGCAGGTTTGTCTGATGTGAAGCCCATTTGAAGACATTCAGACAACTCG 1414
Oy 380 LeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGlnSerSerGlnLeuPheAla 399
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1415 CTACAGAGCTCGAGCTCCAGCTGACGTGAAGTGAACCACTTGAGTCTCACAGCTGTTGCC 1474
Oy 400 LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeuLeu 419
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1475 AAGCTGCTCCAGAAAATGACAGACTCGACAGATTTGTCAGAGAAACGCTCAGCTACTG 1534
Oy 420 GlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeuGlnGluIleTyr 439
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1535 CAGGTGATCAAGAAAGACGGAACAGACATGAGTCTTACCCGCTCGACAGAGATCTAC 1594
Oy 440 LysAspLeuTyrAlaTyrAlaIleLeuThr 449
    ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : :
Db 1595 AAGGACTTGTACTAGCAGAGAGTCTTGAGC 1624
;
RESULT 6
US-09-255-392-1
; Sequence 1, Application US/09255392
; Patent No. 6214850
;
GENERAL INFORMATION:
; APPLICANT: Evans, Ronald M.
; APPLICANT: Forman, Barry M.
; TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR
; TITLE OF INVENTION: ACTIVATED RECEPTOR-GAMMA, AND METHODS FOR THE USE THEREOF
; NUMBER OF SEQUENCES: 6
;
CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
; STREET: 444 South Flower Street, Suite 2000
; CITY: Los Angeles
; STATE: CA
; COUNTRY: USA
; ZIP: 90071
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/255,392
```



```

1 STATE: CALIFORNIA
2
3 COUNTRY: USA
4
5 ZIP: 90071
6
7 COMPUTER READABLE FORM:
8
9 MEDIUM TYPE: Floppy disk
10
11 COMPUTER: IBM PC compatible
12
13 OPERATING SYSTEM: PC-DOS/MS-DOS
14
15 SOFTWARE: Patent Release #1.0.
16
17 CURRENT APPLICATION DATA:
18
19 APPLICATION NUMBER: US/08/484,200
20
21 FILING DATE: 07-JUN-1995
22
23 CLASSIFICATION:
24
25 ATTORNEY/AGENT INFORMATION:
26
27 NAME: REITER, STEPHEN E.
28
29 REGISTRATION NUMBER: 31,192
30
31 REFERENCE/DOCKEN NUMBER: P41 9971
32
33 TELECOMMUNICATION INFORMATION:
34
35 TELEPHONE: 619-546-1995
36
37 TELEFAX: 619-546-9392
38
39 INFORMATION FOR SEQ ID NO: 1:
40
41 SEQUENCE CHARACTERISTICS:
42
43 LENGTH: 2005 base pairs
44
45 TYPE: nucleic acid
46
47 STRANDEDNESS: both
48
49 TOPOLOGY: linear
50
51 MOLECULE TYPE: cDNA
52
53 FEATURE:
54
55 NAME/KEY: CDS
56
57 LOCATION: 352..1776
58
59 US-08-484-200-1

```

```

:
:   NUMBER OF SEQUENCES: 6
:   CORRESPONDENCE ADDRESS:
:   ADDRESS: Pretty, Schroeder, Brueggemann & Clark
:   STREET: 444 South Flower Street, Suite 2000
:   CITY: Los Angeles
:   STATE: CA
:   COUNTRY: USA
:   ZIP: 90071
:   COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: Patent Release #1.0, Version #1.25
:   CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/08/477,493
:   FILING DATE: 07-JUN-1995
:   CLASSIFICATION: 514
:
:   ATTORNEY/AGENT INFORMATION:
:   NAME: Keltner, Stephen E.
:   REGISTRATION NUMBER: 31,192
:   REFERENCE/DOCKET NUMBER: P41 9958
:   TELECOMMUNICATION INFORMATION:
:   TELEPHONE: 619-546-1995
:   TELEFAX: 619-546-9392
:   INFORMATION FOR SEQ ID NO: 1:
:   SEQUENCE CHARACTERISTICS:
:   LENGTH: 2005 base pairs
:   TYPE: nucleic acid
:   STRANDEDNESS: both
:   TOPOLOGY: both
:   MOLECULE TYPE: CDNA
:   FEATURE:
:   NAME/KEY: CDS
:   LOCATION: 352..1776
:   US-08-477-493-1
:
:
: Alignment Scores:
: Pred. No.: 3,21e-232 Length: 2005
: Score: 2065.00 Matches: 408
: Percent Similarity: 86.40% Conservative: 5
: Best Local Similarity: 85.36% Mismatches: 3
: Query Match: 58.87% Indels: 62
: Db: 2 Gaps: 1
:
: US-09-931-007a-1 (1-688) x US-08-477-493-1 (1-2005)
:
: QY 28 GluMetThrMetValAspThrGluMetProPheThrAspPheGlyIleSerSer 47
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 343 GAAATTACCATGTTGACACACAGATGCCATTCTGCCACCACTTCGGAAATCAGCTCT 402
:
: QY 48 ValAspLeuSerValMetGluAspHisSerHisSerPheAspIleLysProPheThrThr 67
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 403 GTGACGCTCTCCGTGATGGAAGACACACGCGCATTCCTTGACATCAAGCCCTTACCACA 462
:
: QY 68 ValAspPheSerSerIleSerThrProHisTyrGluAspIleProPheThrArgThrAsp 87
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 463 GTTGAATTTCTCCAGCATTTCTGCTCCACACATGAAGACATTCCTCAAGACCTGAC 522
:
: QY 88 ProValValAlaAspTyrLysTyrAspLeuLysLeuGlnGluTyrGlnSerAlaIleLys 107
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 523 CCAATGGTGTCTGATTACAAATATGACCTGAAGCTCCAAGAAATACCAAGTGCATCAAA 582
:
: QY 108 ValGluProAlaSerProProTyrTyrSerGluLysThrGlnLeuTyrAsn----- 124
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 583 GTAGAACCCTGCATCTCCACCTTATATCTGAAAGACCCAGGCTGTACAAACAGCGCTCAT 642
:
: QY 124 ----- 124
:
: Db 643 GAAGAACCTTCTAACTCCCTCATGAGCCATGAGTCCGAGTCTGGGGATAAAGCATCA 702
:
: QY 124 ----- 124
:
: Db 703 GGCTTCCACTATGAGTTCATGCTTGTGAGAGGATGCAAGGGTTTTTCCGAAGAACCATC 762
```

```

:
: QY 125 -----Arg 125
:
: Db 763 CGATTGAAGCTTATTATGATAGTGTGATCTTAACCTCCGGATCCAAAAAAGTGA 822
:
: QY 126 AsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAla 145
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 823 AATAATGTAGTACTGTGCTGAGAAAGTCCCTTGCTGTGGGATGCTCAACATGCC 882
:
: QY 146 IleArgPheGlyArgMetProGlnIleGluLysGluLysLeuAlaGluIleSerSer 165
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 883 ATCAGGTTGGGGCGGATGCCAGAGCCGAGAAAGAAAGGAGGCTGTGGCCAGATCTCCAGT 942
:
: QY 166 AspIleAspGlnLeuAsnProGluSerAlaAspLeuArgAlaLeuAlaLysHisLeuTyr 185
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 943 GATATCGACACAGCTGAACCCAGCTCTGATCTGCCAGCCCTTGCAAGACATTTGTAT 1002
:
: QY 186 AspSerTyrIleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGly 205
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 1003 GACTCATACATAAAGTCCCTCCCGCTGACCAAGCCAGAGCGAGGCGATCTTGACAGGA 1062
:
: QY 206 LysThrThrAspLysSerProPheValIleTyrAspMetAsnSerLeuMetMetGlyGlu 225
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 1063 AAGACACGACAAATVCAACCATTTGTCTATCAGCATGAATTCCTTATATGATGGAGAAA 1122
:
: QY 226 AspLysIleLysPheLysHisIleThrProLeuGlnGlnSerLysGluValAlaIle 245
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 1123 GATAAATCAAGTTCAACATATACACCCCTCGCAGAGCAGAGCAAGAGAGTGCGCCATC 1182
:
: QY 246 ArgIlePheGlnGlyCysGlnPheArgSerValGluAlaValGlnGluIleThrGluTyr 265
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 1183 CGAATTTTCAAGGGGTGCAGATTGCGATCCGTAGAAAGCCGTGCAAGATCACAGATAT 1242
:
: QY 266 AlaLysSerIleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLys 285
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 1243 GCCAAAATATCCCGTGTTCATTATTAACCTGATTGAATGACCAAGTACCTGCTCAG 1302
:
: QY 286 TyrGlyValHisGlnIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyVal 305
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 1303 TATGGTGTCCATGAGATATATACACGATCGCTCGCTCCGTGATGAATAAAGATGAGTGC 1362
:
: QY 306 LeuIleSerGluGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysPro 325
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 1363 CTCATCTCAGAGGGCCAGAGATTCATGACACAGGAGGAGTTCCTCAAAAGCCTCGGAATTA 1422
:
: QY 326 PheGlyAspPheMetGluProLysPheGlnPheAlaValLysPheAsnAlaLeuGlnLeu 345
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 1423 TTGGTGACTTATGAGCCCTTAAGTTGAGTTGCTGTGAAGTTCATGACTGGAATTA 1482
:
: QY 346 AspAspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGly 365
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 1483 GATGACATGACTTGGCTATATTTATACCTGTCAATATTCTCAGTGGAGACCGCCAGGC 1542
:
: QY 366 LeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeuLeuGlnLeu 385
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 1543 TTGGTGAACGTGAAGCCCATGAGAGACATCCAAAGACACCTGTCGAGGCCCTGGAATCTG 1602
:
: QY 386 GlnLeuLysLeuAsnHisProGluSerGlnLeuPheAlaLysLeuGlnLysMet 405
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 1603 CAGCTCAAGCTGAATCACCCAGAGCTCTCTCAGCTGTTCGCAAGGCTCTCCAAAGATG 1662
:
: QY 406 ThrAspLeuArgGlnIleValThrGlnHisValGlnLeuLeuGlnValIleLysLysThr 425
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 1663 ACAGACCTCAGACGAGATGCTGACAGAGACGTGACGACTGACATGTGATCAAGAAGACA 1722
:
: QY 426 GluThrAspMetSerLeuHisProLeuLeuGlnGluIleTyrLysAspLeuTyr 443
:   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
: Db 1723 GAGACAGACATGAGCCTTACCCCTGCTCCAGAGATCTCAAGGACTGTAT 1776
:
: RESULT 9
: US-08-465-375-1
: Sequence 1, Application US/08465375A
: Patent No. 6022897
```



```
; TYPE: DNA
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (352)..(1776)
US-09-788-070-1

Alignment Scores:
Pred. No.: 3, 21e-232 Length: 2005
Score: 2065.00 Matches: 408
Percent Similarity: 86.40% Conservative: 5
Best Local Similarity: 85.36% Mismatches: 3
Query Match: 58.87% Indels: 62
DB: Gaps: 1

US-09-931-007a-1 (1-688) x US-09-788-070-1 (1-2005)

QY 28 GUMetThmEValAspHrGluMetProPhetRProThraSnbheGlyIleSer 47
   |||:|||||
Db 343 GAATTTACCATGGTTGACACAGAGATGCCATTCGGCCCCCAACTTCGGAATCAGCTCT 402

QY 48 ValAspLeuSerValMetGluAspHisSerHisSerPheAspIleLysProPhetRThr 67
   |||:|||||
Db 403 GTGGACCTCTCCGTGATGGAAGACCACTGCCATTCCTTGACATCAAGCCCTTTACACA 462

QY 68 ValAspPheSerSerIleSerThrProHisTyrGluAspIleProPhetRThrArgThrAsp 87
   |||:|||||
Db 463 GTTGATTTTCCAGCATTTCTGCTCCACACTATGAAGACATTCATTCACAAAGAGCTGAC 522

QY 88 ProValValAlaAspTyrIleTyrAspLeuLysLeuGlnGluTyrGlnSerAlaIleLys 107
   |||:|||||
Db 523 CCATGGTTCGATTAACAAATATGACCTGAACCTCCAAAGAAATCCAAAGTGCATCAAA 582

QY 108 ValGluProAlaSerProProTyrTyrSerGluLysThrGlnLeuTyrAsn----- 124
   |||:|||||
Db 583 GTGAAACCTGCATCTCCACTTATTATTCTGAAGAACCAAGCTCTACACAGAGCTCAT 642

QY 124 ----- 124
Db 643 GAAGAACCTTCTAATCCCTCATGGCCATGAGTGGCGAGTGTGTGGGGATAAAGCATCA 702

QY 124 ----- 124
Db 703 GGGTTCACATGAGAGTTCAGTCTGTGAAGATGCAAGGGTTTTCGGAAGAACCATC 762

QY 125 -----Arg 125
Db 763 CGATTGAACCTTATTATGATAGGTGTGATCTTAAGTCCGGGATCCACAAAAAAGTATGA 822

QY 126 AsnLysCysGlnTyrIleCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAla 145
   |||:|||||
Db 823 AATTAATGTCAGTACTGTGGTTGCAAGGTGCTTGTGTGGGATGTCTCATCATGCC 882

QY 146 IleArgPheGlyArgMetProGlnAlaGluLysGluLysLeuLeuAlaGluIleSerSer 165
   |||:|||||
Db 883 ATCAGGTTTGGGGGATGGCACAGGCCGGAAGGAAGAAACCTGTGGCGAGATCTCCAGT 942

QY 166 AspIleAspGlnLeuAsnProGlnUserAlaAspLeuArgAlaLeuAlaLysHisLeuTyr 185
   |||:|||||
Db 943 GATATCGACACAGTGAACCAAGATGTGCTGATCTCGAGCCCTGGCAAGCAAGCATTTGTAT 1002

QY 186 AspSerTyrIleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGly 205
   |||:|||||
Db 1003 GACTCATACATTAAGTCTCCCGCTGACCAAAAGCCAAAGCGGCGATCTTGACAGAGA 1062

QY 206 LysThrThrAspLysSerProPheValIleTyrAspMetAsnSerLeuMetLeuTyrGlu 225
   |||:|||||
Db 1063 AAGACACACGACAAATACCATTTTGTCACTACAGCATGAATTTCTTAATGATGGAGAA 1122

QY 226 AspLysIleLysPheLysHisIleThrProLeuGlnGlnGlnSerLysGluValAlaIle 245
   |||:|||||
Db 1123 GATAAATAATCAAGTCAAAATATACACCCCTGCGAGAGACAGCAAAAGAGTGGCCATC 1182
```

```
QY 246 ArgIlePheGlnGlyCysGlnPheArgSerValGluAlaValGlnGluIleThrGluTyr 265
   |||:|||||
Db 1183 CGAATTTTCAAGGTCGCCAGTTTCATCCGTGAAGCCGTGCAGAGATCACAGAGTAT 1242

QY 266 AlaLysSerIleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLys 285
   |||:|||||
Db 1243 GCCAAATATCCCTGGTTTCATTAACCTTGATTTGATGATCCAGAGACTCTGCTCAAG 1302

QY 286 TyrGluValHisGluIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyVal 305
   |||:|||||
Db 1303 TATGTGTCCATAGATCATCTACAGATGCTGGCTCCCTGATGATAAAGATGAGATC 1362

QY 306 LeuIleSerGlnGlyGlnGlyPheMetThrArgGluPheLeuLysSerLeuAaGlyPro 325
   |||:|||||
Db 1363 CTCATCTCAGAGGCCAAGGATTCATGACAGGAGTCTCTCAAAAGCCCTGGCGAAGCCC 1422

QY 326 PheGlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeu 345
   |||:|||||
Db 1423 TTTGGTGACTTGTATGAGAGCTTAAGTTTGAGTTTGTGTGAAGTTCAATGACATCGAATTA 1482

QY 346 AspAspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGly 365
   |||:|||||
Db 1483 GATGACAGTACTTGGCTATATTATTAAGCTGTCTATTATTCATGAGACCCGCCAGGCT 1542

QY 366 LeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGluLeu 385
   |||:|||||
Db 1543 TTCTGACACGTGAAGCCCATCGAGGACATCCAGACACAACTGCTGCAAGGCCCTGGAACCTG 1602

QY 386 GlnLeuLysLeuAsnHisProGlnUserSerGlnLeuPheAlaLysLeuLeuGlnLysMet 405
   |||:|||||
Db 1603 CACCTGAGGTGAATCACCCAGAGTCTCTCACCTGTTCGCCAAAGGTGCTCCAGAAAGATG 1662

QY 406 ThrAspLeuArgGlnIleValThrGluHisValGlnLeuLeuGlnValIleLysLysThr 425
   |||:|||||
Db 1663 ACAGACCTCAGGACAGATCGTCACAGACAGCTGACACTGATGATGATCAAGAGACA 1722

QY 426 GluThrAspMetSerLeuHisProLeuLeuGlnGluIleTyrLysAspLeuTyr 443
   |||:|||||
Db 1723 GAGACAGACATGAGCCCTTCAACCCCTGCTCCAGAGATCTCAAGAGACTTGTAT 1776

RESULT 11
US-08-484-200-3
; Sequence 3, Application US/08484200
; Patent No. 5861274
; GENERAL INFORMATION:
; APPLICANT: EVANS, RONALD M.
; APPLICANT: FORMAN, BARRY M.
; APPLICANT: KLEWER, STEVEN A.
; APPLICANT: ONG, ESTELITA S.
; TITLE OF INVENTION: NOVEL MEMBERS OF THE STEROID/THYROID
; TITLE OF INVENTION: SUPERFAMILY AND USES THEREFOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTY, SCHROEDER, BRUEGEMANN & CLARK
; STREET: 444 SOUTH FLOWER STREET, SUITE 2000
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/484,200
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: REITER, STEPHEN E.
; REGISTRATION NUMBER: 31,192
; REFERENCE/DOCKET NUMBER: P41 9971
; TELECOMMUNICATION INFORMATION:
```



```

Oy 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 665 CTGGCATGTCACACAAAGCCTATCCGCTTGGACGAGATCCGAGCGGCAAGAGAGAG 724
Oy 159 LeuLeuAlaGluIle---SerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeu 177
    ::::::::::::::::::::
Db 725 CTGGGGCGGGGCTACTGCTCCAGCGGGGTCAGACACACCCCAAGCTGACGACCTG 784
Oy 178 ArgAlaLeuAlaLysHisLeuTyraAspSerTyrlleLysSerPheProLeuThrLysAla 197
    ::::::::::::::::::::
Db 785 AAGGCTTCTCTACAGCACATCTACAAAGCCTTACCTGAAAACCTCAACATGACCAAAAG 844
Oy 198 LysAlaAlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyraSp 217
    ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 845 AAGGCGCGGAGATCCTCCACCGGCAAGTCCAGCACCAACGACACCTTTGCTCATCCAGCAG 904
Oy 218 MetAsnSerLeuMetMetLysGluAspLysIleLysPheLysHisIleThrProLeuGln 237
    ::::::::::::::::::::
Db 905 ATCGAGACACTGTGGCAGGACAGAAAGGCGCTGTGTGAACAGCTGTGAACGCGCG 964
Oy 238 GluGlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGlu 257
    ::::::::::::::::::::
Db 965 CCTTACAAAC---GAGATCAGTGTGACAGCTGTTCTACCGCTGCCAGTCCACCAAGTGGAG 1021
Oy 258 AlaValGlnGluIleThrGluTyraLysSerIleProGlyPheValAsnLeuAspLeu 277
    ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1022 ACAGTCCGAGAGCTACCGCAGGATTCGCGCAAGACATCCCACTTCAGCAGCCTCTCTC 1081
Oy 278 AsnAspGlnValThrLeuLeuLysTyrglyValHisGluIleIleTyrrhMetLeuAla 297
    ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1082 AATGACCAAGTACCTCTCAAGATGCGGTGCAGCAGGACGACCTTCCATCGTGGGCC 1141
Oy 298 SerLeuMetAsnLysAspLysValLeuIleSerGluGlyPheMetThrArgGlu 317
    ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1142 TCCATCTGTCAACAAGACGGCGGTGCGTGGCCACAGCGAGTGGCTTCTTCCACCGCAG 1201
Oy 318 PheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAla 337
    ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1202 TTTCTTGGCAAGTCTCCGCAAGCCCTTCACTGACATCATGAGCCCAAGTTGAGTTTGGCT 1261
Oy 338 ValLysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIle 357
    ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1262 GTCAAGTTCAATGCGCTGAGGTCTCGATGACATGACCTGGCGCTTCTTCAATCGGGCCATC 1321
Oy 358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAsp 377
    ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1322 ATTCTGTGTGGAGACCGCGGACGCTCATGATGTGCCCGAGGTAGAACCATCCAGGAC 1381
Oy 378 AsnLeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeu 397
    ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1382 ACCATCTCTCGGGCTCTAGAAATTCCATCTGCAAGTCAACCCCTGACAGCCAGTACCTC 1441
Oy 398 PheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGln 417
    ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1442 TTTCCCAACGCTGCTGAGAAATGCGACCTGCGGACAGTGTGCTCATGAGATGCGCCAG 1501
Oy 418 LeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGlu 437
    ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1502 ATGATGCACTGGCTAAGAGACGAGAGAGTGAACCTTGTGACACCCCTGCTCCAGGAA 1561
Oy 438 IleTyrlLysAspLeuTyra 443
    ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1562 ATCTACAAGCATGTATC 1579

```

```

? APPLICANT: BORGMEYER Ph.D., UWE K.
? APPLICANT: GIGUERE Ph.D., VINCENT NMN
? APPLICANT: YAO M., TSO-PANG NMN
? TITLE OF INVENTION: NOVEL RECEPTORS
? NUMBER OF SEQUENCES: 14
? CORRESPONDENCE ADDRESS:
? ADDRESS: Pretty, Schroeder, Brueggemann & Clark
? STREET: 444 So. Flower St., Suite 2000
? CITY: Los Angeles
? STATE: CA
? COUNTRY: US
? ZIP: 90071-2921
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/694,501
? FILING DATE: 07-AUG-1996
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/08/333,358
? FILING DATE:
? APPLICATION NUMBER: US/07/761,068
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Reiter Ph.D., Stephen E.
? REGISTRATION NUMBER: 31192
? REFERENCE/DOCKET NUMBER: P31 8936
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619) 535-9001
? TELEFAX: (619) 535-8949
? INFORMATION FOR SEQ ID NO: 9:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2009 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? IMMEDIATE SOURCE:
? CLONE: XR4 (XR4.SBG)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 263..1582
? US-08-694-501-9
?
? Alignment Scores:
? Pred. No.: 1,256-115 Length: 2009
? Score: 1071.00 Matches: 202
? Percent Similarity: 80.37% Conservative: 60
? Best Local Similarity: 61.96% Mismatches: 62
? Query Match: 30.53% Indels: 2
? Gaps: 2
?
? US-09-931-007a-1 (1-688) x US-08-694-501-9 (1-2009)
Oy 119 LysThrGlnLeuTyraAsnArgAsnLysCysGlnTyrglyArgPheGlnLysCysLeuAla 138
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 605 AAGATCCACAGAGAGAGACCGCAAGTGTCACTGCTGCGCTTCCAGAAAGTCCGTGGCA 664
Oy 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158
    ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 665 CTGGCATGTCACACACGCTATCCGCTTGGACGAGATCCGAGCGGCAAGAGAGAG 724
Oy 159 LeuLeuAlaGluIle---SerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeu 177
    ::::::::::::::::::::
Db 725 CTGGGGCGGGGCTACTGCTCCAGCGGGGTCAGACACACCCCAAGCTGACGACCTG 784
Oy 178 ArgAlaLeuAlaLysHisLeuTyraAspSerTyrlleLysSerPheProLeuThrLysAla 197
    ::::::::::::::::::::
Db 785 AAGGCTTCTCTACAGCACATCTACAAAGCCTTACCTGAAAACCTCAACATGACCAAAAG 844

```


Db	526	ATGCCAAGATCTGAGAAAGCAAAACGAAAGCAAGAAATTCCTTAAGCTGACACATGACATA	585
Oy	170	LeuasnProgluSerIalaaspleuAargIalaLeuAlaLysHisLeuTyAspSerTyIle	189
Db	586	GAAATCTTGTCAAACTGCAAGTCTCAATCTCTGGCCAAAGAAATCTACAGAGCCCTTCTG	645
Oy	190	LysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLysThrThrasp	209
Db	646	AAGAACTTCACATGACAAACAGGTCCAAAGCCCGGTCATCTCTTCAGGAAAGCCAGTAC	705
Oy	210	LysSerProPheValIleTyraSPmetAsnSerLeuMetMetGlyLusAspLysIleLys	229
Db	706	AATCCACCTTTTGTGCATACATGATATGAGACACTGTATGTGGTGAAGAACGCTGTG	765
Oy	230	PheLysHisIleThrProLeuGlnGluGlnSerLysGluValAlaIleArgIlePheIle	249
Db	766	GCCAAAG--CTGGTGCCCAATGGCATCCAGAACACAGGAGGCGGAGCTCGCATCTTCTAC	822
Oy	250	GlyCysGlnPheArgSerValGluAlaValGlnGluIleThrGlyTyraLysSerIle	269
Db	823	TGCGTCAGTGCACAGTCAGTGAAGTGCACCGCTACGGAGCTCAGGAATTCGCCAAGCCATC	882
Oy	270	ProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysTyGlyValHis	289
Db	883	CCAGGCTTCGGAAACCTTGAGCACTGAACGATCAACGTACATTCCTAAATACGAGATTAT	942
Oy	290	GluIleIleTyrrThrMetLeuAlaSerLeuMetAsnLysAspGlyValIleLysSerL	309
Db	943	GAGGCCATATTCGCCATGCTGCTCTTCTTGATGATACAAAGACGGATGCTGGTACGAT	1002
Oy	310	GlyGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysProPheGlyAspPhe	329
Db	1003	GGAATGAGGTTAATACTCGTAGATTCCTAAAGCCCTAAGAAACCGTCTGGATATTC	1062
Oy	330	MetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeuAspSerAsp	349
Db	1063	ATGGAACCCCAAGTTGATTTGGCCATGAAGTTCAAAGCACTGGAACTGGATGACAGTAT	1122
Oy	350	LeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGlyLeuLeuAsnVal	369
Db	1123	ATTCGCCCTTTTGGCGTGTATCATTTGCTGTGGAGATCGTCTGGCGCTCTAAACSTA	1182
Oy	370	LysProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGluLeuGlnLeuLysLeu	389
Db	1183	GGACACATGTGAAAAAATGCGAGGAGGTATGTACATGTGCTCAGACTCCACCTGACGAGC	1242
Oy	390	AsnHisProGluSerSerGlnLeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArg	409
Db	1243	AACCAACCGGACGATATCTTCTTCCCAAAACTTCTTCAAAAAATGCGACACCTCCGG	1302
Oy	410	GlnIleValThrGlnHisValGlnLeuLeuGlnValIleLysLysThrGluThrAspMet	429
Db	1303	CAGCTGGAGCGAGCATCGCCAGCTGGTGGTGCAGATCATCAAGACAGACGAGCGAGTGT	1362
Oy	430	SerLeuHisProLeuGlnGlnIleTyrrLysAspLeuTyrr	443
Db	1363	GGCTGTCAACCGGTACTGACGAGGAGTCTACAGGAGCATTTAC	1404

Search completed: February 25, 2003, 04:10:34
Job time : 87 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2003, 02:38:25 ; Search time 61 Seconds

(without alignments)
1502.892 Million cell updates/sec

Title: US-09-931-007a-1

Perfect score: 3508
Sequence: 1 MERTLGSDPIPESDSFTDT.....KETDMSLPHLLQETKDLX 688

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3467	98.8	750	22	Human PPARgamma2ga
2	3467	98.8	750	23	Human peroxisome p
3	2351	67.0	467	22	Human PPARgamma2ga
4	2231	63.6	505	17	Peroxisome Prolife
5	2231	63.6	505	22	Human peroxisome p
6	2231	63.6	505	22	Human PPARgamma pr
7	2231	63.6	505	22	Human PPARgamma
8	2106.5	60.0	811	22	Human PAX8e9(-exon
9	2106.5	60.0	874	22	Human PAX8e9-PPARg
10	2099	59.8	777	22	Human PAX8e7-PPARg

11	2094	59.7	494	17	AA99323	Peroxisome prolife
12	2094	59.7	840	22	AA85794	Human PAX8e7-PPARg
13	2089	59.5	477	17	AA99324	Peroxisome prolife
14	2089	59.5	477	17	AA99327	Peroxisome prolife
15	2089	59.5	477	17	AAE1267	Human peroxisome p
16	2079	59.3	475	17	AA99325	Peroxisome prolife
17	2079	59.3	475	17	AA99328	Peroxisome prolife
18	2054.5	58.6	478	20	AA905471	Human PPAR-gamma1
19	2054.5	58.6	478	22	AA85802	Human PPARgamma pr
20	2054.5	58.6	478	22	AA858350	Human peroxisome p
21	2054.5	58.6	478	22	AAE00912	Human peroxisome p
22	2054	58.6	475	17	AA905777	Human peroxisome p
23	2054	58.6	475	17	AA92478	Human peroxisome p
24	2054	58.6	475	20	AAW9596	Peroxisome prolife
25	2044.5	58.3	476	22	AAW9596	Mouse peroxisome p
26	1961	55.9	475	20	AA921635	Human peroxisome p
27	1569	44.7	313	22	AA974068	Ligand binding dom
28	1108.5	31.6	420	20	AA921634	Human colon cancer
29	1101.5	31.4	441	14	AA841875	Steroid hormone re
30	1101.5	31.4	441	17	AA898214	Peroxisome prolife
31	1101.5	31.4	441	20	AA905472	Human PPAR-delta p
32	1076.5	30.7	440	17	AA92479	Peroxisome prolife
33	1076.5	30.7	440	20	AAW9597	Mouse peroxisome p
34	1076.5	30.7	468	16	AA974053	Human peroxisome p
35	1076.5	30.7	468	22	AA920342	Peroxisome prolife
36	1074.5	30.6	468	20	AA921633	Ligand binding dom
37	1074.5	30.6	468	20	AA905470	Human PPAR-alpha p
38	1056	30.1	439	14	AA933745	XR4. Mus musculus
39	524	14.9	121	21	AA853489	Human colon cancer
40	512.5	14.6	128	21	AA853442	Human colon cancer
41	404.5	11.5	579	16	AA971565	EAR-1r gene produc
42	392	11.2	2065	22	AA863705	Drosophila melanog
43	383.5	10.9	1237	12	AA913791	E75A protein. Dro
44	343	9.8	777	22	AA81894	Bovine PPAR. Bos
45	342.5	9.8	548	14	AA935741	XRL. Homo sapiens

ALIGNMENTS

RESULT 1	
AA876987	
ID	AA876987 standard; Protein: 750 AA.
XX	
AC	AA876987;
XX	
DT	22-JUL-2002 (first entry)
XX	
DE	Human PPARgamma2gamma2.
XX	
KW	PPAR response element; PPAR: vaccine; gene therapy; human;
XX	
KW	peroxisome proliferator-activated receptor.
XX	
OS	Homo sapiens.
XX	
PN	WO20078986-A1.
XX	
PD	28-DEC-2000.
XX	
PF	22-JUN-2000; 2000WO-FR01744.
XX	
PR	22-JUN-1999; 99FR-0007957.
XX	
PR	20-AUG-1999; 99US-0149721.
XX	
PA	(AVENT) AVENTIS PHARMA SA.
XX	
PI	Datell R, Crouzet J, Steals B, Mahfoudi A;
XX	
DR	WPI: 2001-091574/10.
XX	
PT	Composition providing inducible expression of a nucleic acid, useful in
PT	gene therapy, uses minimal promoter with peroxisome
PT	proliferator-activated receptor response elements

```

XX Example 5; Page 34; 94pp; French.
PS
CC The present invention relates to a composition (A) comprising a component
CC (A1) containing a nucleic acid (I) controlled by an inducible promoter
CC that consists of a PPAR (peroxisome proliferator-activated receptor)
CC response element (AB158055) and a minimal promoter; and/or a component
CC (A2) comprising a nucleic acid encoding a PPAR under control of a
CC transcriptional promoter. (A), and vectors containing (A1) and (A2), are
CC used to express (I) in cells for expression of transgenic (I) for
CC experimental, clinical, therapeutic or diagnostic purposes. (I) encodes
CC an agriculturally useful, therapeutic, vaccinating or marker protein and
CC is most especially expressed in human muscle cells. Cells containing (A),
CC or the vectors, are used to identify PPAR ligands or to produce
CC transgenic animals for preclinical studies, analysis of bioavailability,
CC labelling etc. The present sequence is human PPAAGamma2, which was
CC used in an example from the invention.
XX
SQ Sequence 750 AA;

Query Match 98.8%; Score 3467; DB 22; Length 750;
Best Local Similarity 91.7%; Pred. No. 1.2e-264;
Matches 688; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 1 MGETLGDSPIDPSDFTDLTANISOEMTMVDTMPFMPPTNGISSVDLSVMEHSHSF 60
DB 1 MGETLGDSPIDPSDFTDLTANISOEMTMVDTMPFMPPTNGISSVDLSVMEHSHSF 60
QY 61 DIPFTVDSSTSTPYEDIPRTDPVADKYDKLQEOYSAIKVEPASPYYSEKT 120
DB 61 DIPFTVDSSTSTPYEDIPRTDPVADKYDKLQEOYSAIKVEPASPYYSEKT 120
QY 121 QLYN----- 124
DB 121 QLYN----- 124
QY 125 -----RNKCOYCFQKCLAVGMSHNAIRGRMPQAEKELAEISSDIDOLNPEADLR 178
DB 181 RIHKRSNKCOCYCRFOKCLAVGMSHNAIRGRMPQAEKELAEISSDIDOLNPEADLR 240
QY 179 ALAKHLVDSTYKSPPLTKAKARAILTGKTDKSPFYVDNNSLMGSDKIKFHIPTLQE 238
DB 241 ALAKHLVDSTYKSPPLTKAKARAILTGKTDKSPFYVDNNSLMGSDKIKFHIPTLQE 300
QY 239 QSKREVAIRIQGCGFRSVEAVOEITEYAKSIPGFVNLDLNDQYTLKYGVEHIIYTMAS 298
DB 301 QSKREVAIRIQGCGFRSVEAVOEITEYAKSIPGFVNLDLNDQYTLKYGVEHIIYTMAS 360
QY 299 LMKKDVLISEGGCFMTRREFLKSIRKPFGEFMEKFEFAVKFNALDELDDSLAIFAVII 358
DB 361 LMKKDVLISEGGCFMTRREFLKSIRKPFGEFMEKFEFAVKFNALDELDDSLAIFAVII 420
QY 359 LSGDRGLNWKRFIEDODNLQALQELQTLNHPSSQLFAKLQKTDLRQIVTEHVOL 418
DB 421 LSGDRGLNWKRFIEDODNLQALQELQTLNHPSSQLFAKLQKTDLRQIVTEHVOL 480
QY 419 LQVTKTETDMSLHPLQELQYKDYAMAILTGKTTDKSPFYVDNNSLMGSDKIKFHHI 478
DB 481 LQVTKTETDMSLHPLQELQYKDYAMAILTGKTTDKSPFYVDNNSLMGSDKIKFHHI 540
QY 479 TPLQDSQKEVAIRIQGCGFRSVEAVOEITEYAKSIPGFVNLDLNDQYTLKYGVEHII 538
DB 541 TPLQDSQKEVAIRIQGCGFRSVEAVOEITEYAKSIPGFVNLDLNDQYTLKYGVEHII 600
QY 539 TMLASLNNKGVLISSGGCFMTRREFLKSIRKPFGEFMEKFEFAVKFNALDELDDSLAIF 558
DB 601 TMLASLNNKGVLISSGGCFMTRREFLKSIRKPFGEFMEKFEFAVKFNALDELDDSLAIF 660
QY 599 IAVIIISGDRPGLNWKRFIEDODNLQALQELQTLNHPSSQLFAKLQKTDLRQIVT 658
DB 661 IAVIIISGDRPGLNWKRFIEDODNLQALQELQTLNHPSSQLFAKLQKTDLRQIVT 720
QY 659 EHVOLLQVIRKKTETDMSLHPLQELQYKDY 688

```

```

DB 721 EHVOLLQVIRKKTETDMSLHPLQELQYKDY 750
|||||
RESULT 2
ABB09779
ID ABB09779 standard; protein; 750 AA.
XX
AC ABB09779;
XX
DT 22-JUL-2002 (first entry)
XX
DE Human peroxidome proliferator activated receptor (PPAR)alpha.
XX
KW Human; peroxidome proliferator activated receptor alpha; PPARalpha;
KW transgene; mitochondrial disease; myopathy; ischaemia; stenosis;
KW lysosomal storage disease; hormonal disorder; haemophilia; inflammation;
KW rheumatoid polyarthritits; beta-thalassemia; cancer; neurodegeneration;
KW cardiovascular disease; hypertension; hyperlipidaemia; obesity;
KW receptor; vaccine.
XX
XX Homo sapiens.
OS
XX
XX WO200213758-A2.
XX
XX 21-FEB-2002.
XX
XX 10-AUG-2001; 2001WO-FR02606.
XX
XX 18-ANG-2000; 2000FR-0010730.
XX
XX 11-OCT-2000; 2000US-239246P.
XX
XX (AVENTIS PHARMA SA.
XX
XX Scherman D, Bettan M, Bigey P;
XX
XX WPI: 2002-269145/31.
XX
XX Regulating expression of transgenes in plants and animals; useful e.g.
XX PT for gene therapy, comprises cotransfection with a transgene and a
XX PT sequence that expresses an inhibitory transcript
XX
XX PS Disclosure; Page 15; 123pp; French.
XX
XX The present sequence represents human peroxidome proliferator activated
XX receptor (PPAR)alpha. This receptor may be used to regulate expression
XX of a transgene. The specification describes a method for regulating
XX expression of a selected transgene in vivo. The method comprises
XX introducing, into a non-human animal tissue or target cell a nucleic
XX acid comprising the transgene and encoding a transcript (T1); and
XX a nucleic acid that encodes a transcript (T2) that inhibits T1
XX specifically. Both nucleic acids are co-expressed so that activity of
XX T1 is inhibited constitutively by T2. The nucleic acids are controlled
XX by a transcriptional promoter and the activity of T2 and/or T1 can be
XX regulated by an external agent. The method is used to regulate a
XX transgene, in animals and plants, particularly for control of therapeutic
XX transgenes in treatment of genetic anomalies and defects, e.g.
XX mitochondrial diseases, myopathy, ischaemia, stenosis, lysosomal storage
XX diseases, hormonal disorders, haemophilia, inflammation (rheumatoid
XX polyarthritits), beta-thalassemia, cancer (by inducing apoptosis or
XX expression of toxins), neurodegeneration, cardiovascular diseases
XX (hypertension), hyperlipidaemia (obesity), and in preparation of
XX vaccines. Transgenic animals containing T1 and T2 are useful as
XX experimental models of diseases and transgenic plants are useful for
XX studying the effects of specific genes on development etc.
XX
XX Sequence 750 AA;

Query Match 98.8%; Score 3467; DB 23; Length 750;
Best Local Similarity 91.7%; Pred. No. 1.2e-264;
Matches 688; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 1 MGETLGDSPIDPSDFTDLTANISOEMTMVDTMPFMPPTNGISSVDLSVMEHSHSF 60

```

```

Db      1 MGELTLDSPIDPESDFTLSTANISQEMTMVDEMPNPTNGISVDSLWEDSHSP 60
QY      61 DIRPFTTVDFSSISTPHYEDIEDPFRTRDPVAVADKYDLKLOEYSAIKVEBPASPPYSEKT 120
Db      61 DIRPFTTVDFSSISTPHYEDIEDPFRTRDPVAVADKYDLKLOEYSAIKVEBPASPPYSEKT 120
QY      121 QLYN----- 124
Db      121 QLYNKHPEPSNSLMAIECNVCGDKASGFHYGVHACEGCKGFFRRITRLKLIYRCDLNC 180
QY      125 -----RNKCQYCRFOKCLAVGSHNAIRFGMPQAEKEKILAEISSDIDOLNPESADLR 178
Db      181 RIHKSRNKCQYCRFOKCLAVGSHNAIRFGMPQAEKEKILAEISSDIDOLNPESADLR 240
QY      179 ALAKHLYDSYIKSFPLTKAKARAILTGKTTDKSPFYIYDMNSLMGEDIKFKHITPLOE 238
Db      241 ALAKHLYDSYIKSFPLTKAKARAILTGKTTDKSPFYIYDMNSLMGEDIKFKHITPLOE 300
QY      239 QSKREVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLDLNDQVTLTKYGVHEIITMLAS 298
Db      301 QSKREVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLDLNDQVTLTKYGVHEIITMLAS 360
QY      299 LMKKDGVLISEGQGFMTREFLSLRKPGDFMPEKFEFAVKFNALDELDDSLAIFIAVII 358
Db      361 LMKKDGVLISEGQGFMTREFLSLRKPGDFMPEKFEFAVKFNALDELDDSLAIFIAVII 420
QY      359 LSGDRGLNVKPIEDIONLLOALELQKLNHPRESSOLFATLQKMTDLROIYTEHVOL 418
Db      421 LSGDRGLNVKPIEDIONLLOALELQKLNHPRESSOLFATLQKMTDLROIYTEHVOL 480
QY      419 LOYIKKETDMSLHPLOEIYKDYAMAILTGKTTDKSPFYIYDMNSLMGEDIKFKHI 478
Db      481 LOYIKKETDMSLHPLOEIYKDYAMAILTGKTTDKSPFYIYDMNSLMGEDIKFKHI 540
QY      479 TPLOESKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLDLNDQVTLTKYGVHEIIT 538
Db      541 TPLOESKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLDLNDQVTLTKYGVHEIIT 600
QY      539 TMLASLMNKDGVILISEGQGFMTREFLSLRKPGDFMPEKFEFAVKFNALDELDDSLAIF 598
Db      601 TMLASLMNKDGVILISEGQGFMTREFLSLRKPGDFMPEKFEFAVKFNALDELDDSLAIF 660
QY      599 IAVIILSGDRPGLNVKPIEDIONLLOALELQKLNHPRESSOLFATLQKMTDLROIYV 658
Db      661 IAVIILSGDRPGLNVKPIEDIONLLOALELQKLNHPRESSOLFATLQKMTDLROIYV 720
QY      659 EHVOLLQVIAKKTETDMSLHPLOEIYKDY 688
Db      721 EHVOLLQVIAKKTETDMSLHPLOEIYKDY 750

```

RESULT 3

ABB76988 standard; Protein: 467 AA.

ABBT6988:

22-JUL-2002 (first entry)

Human PPARgamma2gamma2 E/F domain.

PPAR response element; PPAR: vaccine; gene therapy: human; peroxisome proliferator-activated receptor.

Homo sapiens.

WO200078986-A1.

28-DEC-2000.

22-JUN-2000; 2000WO-FR01744.

XX

```

PR      22-JUN-1999; 99FR-0007957.
PR      20-AUG-1999; 99US-0149721.
XX      (AVET ) AVENTIS PHARMA SA.
XX      Darrell R, Crouzet J, Staels B, Mahfoudi A;
XX      WPI: 2001-091574/10.
XX
XX      Composition providing inducible expression of a nucleic acid, useful in
XX      gene therapy, uses minimal promoter with peroxisome
XX      proliferator-activated receptor response elements
XX
XX      Example 5; Page 34; 94pp; French.
XX
XX      The present invention relates to a composition (A) comprising a component
XX      (A1) containing a nucleic acid (1) controlled by an inducible promoter
XX      that consists of a PPAR (peroxisome proliferator-activated receptor)
XX      response element (Abl58055) and a minimal promoter; and/or a component
XX      (A2) comprising a nucleic acid encoding a PPAR under control of a
XX      transcriptional promoter. (A), and vectors containing (A1) and (A2), are
XX      used to express (1) in cells for expression of transgenic (1) for
XX      experimental, clinical, therapeutic or diagnostic purposes. (1) encodes
XX      an agriculturally useful, therapeutic, vaccinating or marker protein and
XX      is most especially expressed in human muscle cells. Cells containing (A),
XX      or the vectors, are used to identify PPAR ligands or to produce
XX      transgenic animals for preclinical studies, analysis of bioavailability,
XX      labelling etc. The present sequence is the E/F domain of human
XX      PPARgamma2gamma2, which was used in an example from the invention.
XX
XX      Sequence 467 AA:
XX
XX      Query Match 67 0%; Score 2351; DB 22; Length 467;
XX      Best Local Similarity 100.0%; Pred. No. 5.9e-177;
XX      Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      222 NMGEKIKFKHITPLOESKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLNDQV 281
Db      1 NMGEKIKFKHITPLOESKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLNDQV 60
QY      282 TLTKYGVHEIITMLASLMNKDGVILISEGQGFMTREFLSLRKPGDFMPEKFEFAVKFN 341
Db      61 TLTKYGVHEIITMLASLMNKDGVILISEGQGFMTREFLSLRKPGDFMPEKFEFAVKFN 120
QY      342 ALELDSDLAIFIAVIIISGDRPGLNVKPIEDIONLLOALELQKLNHPRESSOLFATL 401
Db      121 ALELDSDLAIFIAVIIISGDRPGLNVKPIEDIONLLOALELQKLNHPRESSOLFATL 180
QY      402 LQKMTDLROIYTEHVOLLQVIAKKTETDMSLHPLOEIYKDYAMAILTGKTTDKSPFYI 461
Db      181 LQKMTDLROIYTEHVOLLQVIAKKTETDMSLHPLOEIYKDYAMAILTGKTTDKSPFYI 240
QY      462 DMNSLMGEDIKFKHITPLOESKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVND 521
Db      241 DMNSLMGEDIKFKHITPLOESKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVND 300
QY      522 LNDQVTLTKYGVHEIITMLASLMNKDGVILISEGQGFMTREFLSLRKPGDFMPEKFEF 581
Db      301 LNDQVTLTKYGVHEIITMLASLMNKDGVILISEGQGFMTREFLSLRKPGDFMPEKFEF 360
QY      582 AVKFNALELDSDLAIFIAVIIISGDRPGLNVKPIEDIONLLOALELQKLNHPRESSQ 641
Db      361 AVKFNALELDSDLAIFIAVIIISGDRPGLNVKPIEDIONLLOALELQKLNHPRESSQ 420
QY      642 LFAKLLQKMTDLROIYTEHVOLLQVIAKKTETDMSLHPLOEIYKDY 688
Db      421 LFAKLLQKMTDLROIYTEHVOLLQVIAKKTETDMSLHPLOEIYKDY 467

```

RESULT 4

AAR99326 standard; Protein: 505 AA.

XX

```
AC AAR99326;
XX
XX 14-APR-1997 (first entry)
XX
XX Peroxisome proliferator activated receptor gamma2 subtype 1.
DE
XX Peroxisome proliferator activated receptor gamma; hppargamma; human;
XX peroxisome; cholesterol metabolism; lipid metabolism; respiration;
XX fatty acid metabolism; beta-oxidation cycle enzyme; hypolipidemic drug;
XX herbicide; hppargamma2; adipose tissue; obesity; diabetes; anorexia;
XX hyperlipidaemia; hypercholesterolaemia; hyperlipoproteinaemia; antibody;
XX metabolic disorder; therapy.
XX
XX Homo sapiens.
XX
XX MO9623884-A2.
XX
XX 08-AUG-1996.
XX
XX 29-JAN-1996; 96WO-US01469.
XX
XX 23-OCT-1995; 95US-0005809.
XX 30-JAN-1995; 95US-0380051.
XX 07-JUN-1995; 95US-0484487.
XX
XX (LIGA-) LIGAND PHARM INC.
XX
XX Mukherjee R;
XX
XX WPI; 1996-384114/38.
XX N-PSDB; AAT35334.
XX
XX Human peroxisome proliferator activated receptor(s), hppar-gamma and
XX hppar-gamma2 - useful to identify polypeptide (ant)agonists to treat
XX e.g. obesity, anorexia and diabetes
XX
XX Claim 43; Page 89-92; 108pp; English.
XX
XX AAR99326-R99328 represent the three different subtypes of human
XX peroxisome proliferator activated receptor gamma2 (hppargamma2). These
XX sequences differ in their N-terminal sequences. Peroxisomes contain
XX enzymes for cholesterol and lipid metabolism, and respiration.
XX Peroxisome proliferators increase the capacity of the peroxisomes to
XX metabolise fatty acids, via increased expression of the beta-oxidation
XX cycle enzymes. Peroxisome proliferators include unsaturated fatty
XX acids, hypolipidemic drugs, and herbicides. ppargamma has two main
XX subtypes (hppargamma, and hppargamma2), which differ in their N-terminal
XX sequences, and are expressed at high levels in adipose tissue.
XX hppargamma proteins repress hpparalpha activity. These proteins can be
XX used in a cell system with a reporter gene (and optionally a hpparalpha
XX protein) to identify agonists and antagonists of the polypeptides, which
XX can be used as therapeutic agents. The agonists and antagonists can be
XX used to treat obesity, diabetes, anorexia, hyperlipidaemia,
XX hypercholesterolaemia, hyperlipoproteinaemia, and other metabolic
XX disorders. These sequences, the DNA encoding them, and antibodies
XX against them, can be used to establish a tissue specific expression
XX pattern for hppargamma.
XX
XX SQ Sequence 505 AA;
XX
XX Query Match 63.6%; Score 2231; DB 17; Length 505;
XX Best Local Similarity 87.7%; Pred. No. 1.9e-167;
XX Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;
```

```
DB 121 QLYNKPHEEPSNLSMAIECVCGDKASGFHYGVHACEGCKGFFRRRTLRKLIYDRCLNC 180
|||||
DB 125 -----RNKCQYCRFOKCLAVGSMHNAIRGRMPQAEKEKLAIESSDDLDLPESADLR 178
|||||
DB 181 RIHKSRNKCQYCRFOKCLAVGSMHNAIRGRMPQAEKEKLAIESSDDLDLPESADLR 240
|||||
QY 179 ALAKHLYDSYIKSFPLTKAKARAILTGKTTDKSPFIYOMNSLMGMDKIKFKHITPLOE 238
|||||
DB 241 ALKKHLYDSYIKSFPLTKAKARAILTGKTTDKSPFIYIDMNSLMGMDKIKFKHITPLOE 300
|||||
QY 239 QSKKEVAIRIFQCGCFRSVEAVQETFEYAKSIPGFVNLNDLNDQYTLKYGHEIITYMLAS 298
|||||
DB 301 QSKKEVAIRIFQCGCFRSVEAVQETFEYAKSIPGFVNLNDLNDQYTLKYGHEIITYMLAS 360
|||||
QY 299 LMKKGVLSSEGQGFRTREFLKSRLKRPFGDMEPKKEFAVKFNAELDDSDLAIFAVI 358
|||||
DB 361 LMKKGVLSSEGQGFRTREFLKSRLKRPFGDMEPKKEFAVKFNAELDDSDLAIFAVI 420
|||||
QY 359 LSGDRPGLNVKRPTEIDQNLQALELQKLNHPSSQLFAKLQKMTDLROIYTEHVOL 418
|||||
DB 421 LSGDRPGLNVKRPTEIDQNLQALELQKLNHPSSQLFAKLQKMTDLROIYTEHVOL 480
|||||
QY 419 LQYIKKTEYDMSLHPLQEIYKDLV 443
|||||
DB 481 LQYIKKTEYDMSLHPLQEIYKDLV 505
|||||

RESULT 5
AAE12868
ID AAE12868 standard; Protein; 505 AA.
XX
XX AC AAE12868;
XX
XX DT 15-JAN-2002 (first entry)
XX
XX DE Human peroxisome proliferator activated receptor (PPAR) gamma2 protein.
XX
XX KW Human; peroxisome proliferator activated receptor gamma2; PPAR gamma2;
XX cytotoxic; antiproliferative; antiviral; cancer; cell proliferation;
XX viral infection; pharmaceutical; thiazolidinedione.
XX
XX OS Homo sapiens.
XX
XX PN US6294559-B1.
XX
XX PD 25-SEP-2001.
XX
XX PF 03-AUG-1998; 98US-0128142.
XX
XX PR 02-MAY-1996; 96US-016694P.
XX PR 18-APR-1997; 97US-0844007.
XX
XX PA (MERI ) MERCK & CO INC.
XX
XX PI Smith RG;
XX
XX DR WPI; 2001-647265/74.
XX DR N-PSDB; AAD21022.
XX
XX Use of thiazolidinedione for treating cancer and viral infections -
XX
XX Example 3; Fig 4; 17pp; English.
XX
XX CC The invention relates to compounds and ligands that bind to human
XX CC peroxisome proliferator activated receptors (PPAR) gamma1 and gamma2.
XX CC The invention is useful for treating cancer and other disorders including
XX CC excessive cell proliferation and viral infection. The invention is also
XX CC directed to the use of PPAR gamma1 and gamma2 to identify compounds that
XX CC are antiproliferative, antiviral and antitumour agents. The invention
XX CC also relates to a method of treating cancer using a pharmaceutical
XX CC composition comprising thiazolidinedione in an amount sufficient to
XX CC modulate PPAR gamma1 and gamma2 activity. The present sequence is human
```

CC peroxisome proliferator activated receptor (PPAR) gamma2 protein related
CC to the invention.
XX
SQ Sequence 505 AA:
Query Match 63.6%; Score 2231; DB 22; Length 505;
Best Local Similarity 87.7%; Pred. No. 1.9e-167;
Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;
QY 1 MGETLGDSPIDPESDSFTDTLSANISOEMTWDTMPMPPTNFGISSVDLSVMEHSHSF 60
DB 1 MGETLGDSPIDPESDSFTDTLSANISOEMTWDTMPMPPTNFGISSVDLSVMEHSHSF 60
QY 61 DIKPEFTVDFSSISTPHYEDIPFRTDPVADYKYDKLQEQSAIKVEPASPYYSEKT 120
DB 61 DIKPEFTVDFSSISTPHYEDIPFRTDPVADYKYDKLQEQSAIKVEPASPYYSEKT 120
QY 121 QLYN----- 124
DB 121 QLYNRPHEEPSNSLMAIECRVCGDKASGFHYGHACGCKGFRTIRLKLIDRCDLNC 180
QY 125 -----RNKCYCRFOCKLAVGSMHNATRFGRMPOAEKEKLLAETSSDIDOLNPESADLR 178
DB 181 RIHKSRNKCQYCRFOCKLAVGSMHNATRFGRMPOAEKEKLLAETSSDIDOLNPESADLR 240
QY 179 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMNGEDKIKFKHTTLPQE 238
DB 241 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMNGEDKIKFKHTTLPQE 300
QY 239 OSKEVAIRIFOGCQFRSYEAVOEITEYAKSIPGFVNLDNDQVTLKYGVEHIIYTMLAS 298
DB 301 OSKEVAIRIFOGCQFRSYEAVOEITEYAKSIPGFVNLDNDQVTLKYGVEHIIYTMLAS 360
QY 299 LMNKGVLISSEGQGMTEPEFLKSLRPGDFMEPKFEFAVKNFNALEDDSLAIFIAVII 358
DB 361 LMNKGVLISSEGQGMTEPEFLKSLRPGDFMEPKFEFAVKNFNALEDDSLAIFIAVII 420
QY 359 LSGDRPGLLVNKPRIEDIDNLQALELQKLHNPESSQLFAKLLQKMTDLRQIYTEHVQL 418
DB 421 LSGDRPGLLVNKPRIEDIDNLQALELQKLHNPESSQLFAKLLQKMTDLRQIYTEHVQL 480
QY 419 LQVTKTETDMSLHPLLOEIYKDL 443
DB 481 LQVTKTETDMSLHPLLOEIYKDL 505
RESULT 6
AAB5800
ID AAB5800 standard; Protein: 505 AA.
XX
AC AAB5800;
XX
DT 29-OCT-2001 (first entry)
XX
DE Human PPARgamma protein sequence.
XX
KW PAX8-PPARGamma1; oncogene; cytosstatic; PAX8; PPARgamma1; cancer;
XX
OS follicular carcinoma; PPARgamma; human.
XX
XX Homo sapiens.
XX
XX WO200152789-A2.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US01664.
XX
XX 20-JAN-2000; 2000US-0177109.
XX
XX 14-AUG-2000; 2000US-0225079.
XX
XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Kroll TG, Fletcher JA.
XX
XX PI

XX
DR WPI: 2001-514487/56.
XX
DR N-PSDB; AAH76288.
XX
PT New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma
XX
PS Disclosure: Page 123-125; 145pp; English.
XX
CC The invention relates to an oncogene designated PAX8-PPARGamma1 that
CC contains a PAX8 coding region fused to PPARgamma1 coding region. The
CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant
CC methodology. A PPARgamma1 ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARgamma1, where the disorder is cancer, e.g. follicular carcinoma.
CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The
CC present sequence represents a human PPARgamma polypeptide.
XX
SQ Sequence 505 AA:
Query Match 63.6%; Score 2231; DB 22; Length 505;
Best Local Similarity 87.7%; Pred. No. 1.9e-167;
Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;
QY 1 MGETLGDSPIDPESDSFTDTLSANISOEMTWDTMPMPPTNFGISSVDLSVMEHSHSF 60
DB 1 MGETLGDSPIDPESDSFTDTLSANISOEMTWDTMPMPPTNFGISSVDLSVMEHSHSF 60
QY 61 DIKPEFTVDFSSISTPHYEDIPFRTDPVADYKYDKLQEQSAIKVEPASPYYSEKT 120
DB 61 DIKPEFTVDFSSISTPHYEDIPFRTDPVADYKYDKLQEQSAIKVEPASPYYSEKT 120
QY 121 QLYN----- 124
DB 121 QLYNRPHEEPSNSLMAIECRVCGDKASGFHYGHACGCKGFRTIRLKLIDRCDLNC 180
QY 125 -----RNKCYCRFOCKLAVGSMHNATRFGRMPOAEKEKLLAETSSDIDOLNPESADLR 178
DB 181 RIHKSRNKCQYCRFOCKLAVGSMHNATRFGRMPOAEKEKLLAETSSDIDOLNPESADLR 240
QY 179 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMNGEDKIKFKHTTLPQE 238
DB 241 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMNGEDKIKFKHTTLPQE 300
QY 239 OSKEVAIRIFOGCQFRSYEAVOEITEYAKSIPGFVNLDNDQVTLKYGVEHIIYTMLAS 298
DB 301 OSKEVAIRIFOGCQFRSYEAVOEITEYAKSIPGFVNLDNDQVTLKYGVEHIIYTMLAS 360
QY 299 LMNKGVLISSEGQGMTEPEFLKSLRPGDFMEPKFEFAVKNFNALEDDSLAIFIAVII 358
DB 361 LMNKGVLISSEGQGMTEPEFLKSLRPGDFMEPKFEFAVKNFNALEDDSLAIFIAVII 420
QY 359 LSGDRPGLLVNKPRIEDIDNLQALELQKLHNPESSQLFAKLLQKMTDLRQIYTEHVQL 418
DB 421 LSGDRPGLLVNKPRIEDIDNLQALELQKLHNPESSQLFAKLLQKMTDLRQIYTEHVQL 480
QY 419 LQVTKTETDMSLHPLLOEIYKDL 443
DB 481 LQVTKTETDMSLHPLLOEIYKDL 505
RESULT 7
AAB59839
ID AAB59839 standard; Protein: 505 AA.
XX
AC AAB59839;
XX
DT 27-MAR-2001 (first entry)
XX
XX Human PPARgamma.
XX
XX Cytostatic; antiinflammatory; antisense oligonucleotide; PPARgamma;
XX
XX

KW peroxisome proliferator-activated receptor gamma; transcription factor;
 KW nuclear hormone receptor; human; infection; inflammation; tumour.
 XX
 OS Homo sapiens.
 XX US6159734-A.
 XX
 PD 12-DEC-2000.
 XX
 PF 18-JAN-2000; 2000US-0484345.
 XX
 PR 18-JAN-2000; 2000US-0484345.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI McKay R, Baker BF, Borchers AH;
 XX
 DR WPI: 2001-070112/08.
 DR N-PSDB: AAF23644.
 XX
 PT Novel antisense compounds capable of modulating expression of
 PT peroxisome proliferator-activated receptor gamma useful for diagnosis,
 PT prophylaxis and treatment of diseases associated with expression of the
 PT receptor -
 XX
 PS Example 13; Column 45-48; 40pp; English.
 XX
 CC Peroxisome proliferator-activated receptors (PPARs) are members of the
 CC nuclear hormone receptor subfamily of transcription factors. The present
 CC invention relates to antisense oligonucleotides (see AAF23652-F23731),
 CC targeted to a nucleic acid molecule encoding human PPARgamma, which
 CC specifically hybridises with and inhibits the expression of human
 CC PPARgamma. The present sequence is human PPARgamma. The PPARgamma
 CC antisense oligonucleotides of the present invention can be used in the
 CC diagnosis and treatment of diseases associated with the expression of
 CC PPARgamma, e.g. to prevent or delay infection, inflammation or tumour
 CC formation.
 CC
 SQ Sequence 505 AA;
 Query Match 63.6%; Score 2231; DB 22; Length 505;
 Best Local Similarity 87.7%; Pred. No. 1.9e-167;
 Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;
 QY 1 MGETLGDSFIDEDSDSFDTLSANISQEMTWDTMPEWPTNFGISSVDLSVMEHDSHSF 60
 DB 1 MGETLGDSFIDEDSDSFDTLSANISQEMTWDTMPEWPTNFGISSVDLSVMEHDSHSF 60
 QY 61 DIKPTTVDFSSISTPHYEDIPFTRTPVADYKYDKLQEYQSAIKYEPASPYYSEKT 120
 DB 61 DIKPTTVDFSSISTPHYEDIPFTRTPVADYKYDKLQEYQSAIKYEPASPYYSEKT 120
 QY 121 QLYN----- 124
 DB 121 QLYNPRHEEPSNSLMAIECRVCGDKASGFHYGVHACEGCKGFRTIRLKLIVDRCDLNC 180
 QY 125 -----RNKCOYCRFOKCLAVGSMHNAIIRFGMPQAEKEKLLAETSSDIDOLNPESADLR 178
 DB 181 RTHKSSRNKCOYCRFOKCLAVGSMHNAIIRFGMPQAEKEKLLAETSSDIDOLNPESADLR 240
 QY 179 ALAKHLUDSYIKSPFLTKAKARAILGKTGTTDKSPVIYDMNSLMGCEDKIKFKHITPJOE 238
 DB 241 ALAKHLUDSYIKSPFLTKAKARAILGKTGTTDKSPVIYDMNSLMGCEDKIKFKHITPJOE 300
 QY 239 OSKEVAIRIFQCCQFRSVAVOEITEYAKSIFGFVNLINDOVTLTKYGVHEIITTMLAS 298
 DB 301 OSKEVAIRIFQCCQFRSVAVOEITEYAKSIFGFVNLINDOVTLTKYGVHEIITTMLAS 360
 QY 299 LNNKGVLISSEGGFMTREFLTKLRKPCGFMPEKFEFVKNFNALELDDSDLAIFIAVII 358
 DB 361 LNNKGVLISSEGGFMTREFLTKLRKPCGFMPEKFEFVKNFNALELDDSDLAIFIAVII 420
 QY 359 LSGDRGLINVKPIEDIODNLQALELQIKLNHPRESSQLFAKLQKMTDLROIIVTEHVOL 418

DB 421 LSGDRGLINVKPIEDIODNLQALELQIKLNHPRESSQLFAKLQKMTDLROIIVTEHVOL 480
 QY 419 LOVIKKTETDMSLHPLQEIYKLDY 443
 DB 481 LOVIKKTETDMSLHPLQEIYKLDY 505
 RESULT 8
 AAB85801
 ID AAB85801 standard; Protein; 811 AA.
 XX
 AC AAB85801;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Human PAX8e9(-exon 8)-PPARgammae1 protein sequence.
 XX
 KW PAX8-PPARgammae1; oncogene; cytostatic; PAX8; PPARgammae1; cancer;
 KW follicular carcinoma; PAX8e9(-exon 8)-PPARgammae1; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200152789-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001WO-US01664.
 XX
 PR 20-JAN-2000; 2000US-0177109.
 PR 14-AUG-2000; 2000US-0225079.
 XX
 PA (BCHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 PI Kroll TG, Fletcher JA;
 XX
 DR WPI: 2001-514487/56.
 DR N-PSDB: AAH76294.
 XX
 PT New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
 PT treating certain tumors or cancers, e.g. follicular carcinoma -
 XX
 PS Claim 16; Page 127-131; 145pp; English.
 XX
 CC The invention relates to an oncogene designated PAX8-PPARgammae1 that
 CC contains a PAX8 coding region fused to PPARgammae1 coding region. The
 CC PAX8-PPARgammae1 polypeptides can be expressed by standard recombinant
 CC methodology. A PPARgammae1 ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARgammae1, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARgammae1 molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents a human PAX8e9(-exon 8)-PPARgammae1
 CC polypeptide.
 CC
 SQ Sequence 811 AA;
 Query Match 60.0%; Score 2106.5; DB 22; Length 811;
 Best Local Similarity 83.7%; Pred. No. 2.4e-157;
 Matches 425; Conservative 4; Mismatches 14; Indels 65; Gaps 3;
 QY 1 MGETL-GDSPIDPES--DSFTDTLSANISQEMTWDTMPEWPTNFGISSVDLSVMEHDS 57
 DB 304 VGPPLRGYPHPHPTPGSGGSYASSALAGVAEMTWDTMPEWPTNFGISSVDLSVMEHDS 363
 QY 58 HSFEDIKPTTVDFSSISTPHYEDIPFTRTPVADYKYDKLQEYQSAIKYEPASPYYYS 117
 DB 364 HSFEDIKPTTVDFSSISTPHYEDIPFTRTPVADYKYDKLQEYQSAIKYEPASPYYYS 423
 QY 118 EKTOLYXN----- 124
 DB 424 EKTOLYXNPRHEEPSNSLMAIECRVCGDKASGFHYGVHACEGCKGFRTIRLKLIVDRCD 483

QY 125 -----RNCQYCRFOKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDQLPESA 175
 |||
 DB 484 LNCRIHKKSRNKCQYCRFOKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDQLPESA 543
 QY 176 DLRALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFYIDMNSLMGEDKIKFKHITP 235
 |||
 DB 544 DLRALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFYIDMNSLMGEDKIKFKHITP 603
 QY 236 LQOSKEVAIRIFOGQCFRSEVAVOETETAKSIPGVNLDLNDQYTLTKYGVHEIITYM 295
 |||
 DB 604 LQOSKEVAIRIFOGQCFRSEVAVOETETAKSIPGVNLDLNDQYTLTKYGVHEIITYM 663
 QY 296 LASIMNKDGVLISEGQCFMTREFLSLRKPFQDMEPKFEFAVKFNALELDSDLAIFIA 355
 |||
 DB 664 LASIMNKDGVLISEGQCFMTREFLSLRKPFQDMEPKFEFAVKFNALELDSDLAIFIA 723
 QY 356 VIILSGDRPGLNVKPIEDIDQNLQALELQKLNPRESSQFAKLLQKMTDLRQIVTEH 415
 |||
 DB 724 VIILSGDRPGLNVKPIEDIDQNLQALELQKLNPRESSQFAKLLQKMTDLRQIVTEH 783
 QY 416 VQLLQVYIKKTEPDMSLPILQEIYKDL 443
 |||
 DB 784 VQLLQVYIKKTEPDMSLPILQEIYKDL 811
 RESULT 9
 AAB85795
 ID AAB85795 standard; Protein; 874 AA.
 XX
 AC AAB85795;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Human PAX8e9-PPARGammael protein sequence.
 XX
 KW PAX8-PPARGammael; oncogene; cytostatic; PAX8; PPARGammael; cancer;
 XX
 KW Follicular carcinoma; PAX8e9-PPARGammael; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200152789-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001WO-US01664.
 XX
 PR 20-JAN-2000; 2000US-0177109.
 XX
 PR 14-AUG-2000; 2000US-0225079.
 XX
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 PI Kroll TG, Fletcher JA;
 XX
 DR WPI; 2001-514487/56.
 XX
 DR N-PSDB; AAH76283.
 XX
 PT New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
 XX
 PT treating certain tumors or cancers, e.g. follicular carcinoma -
 XX
 PS Claim 16; Page 112-116; 145pp; English.
 XX
 CC The invention relates to an oncogene designated PAX8-PPARGammael that
 XX
 CC contains a PAX8 coding region fused to PPARGammael coding region. The
 CC PAX8-PPARGammael polypeptides can be expressed by standard recombinant
 CC methodology. A PPARGammael ligand or agent is useful for treating a
 CC subject having a disorder characterized by the presence of a PAX8-
 CC PPARGammael, where the disorder is cancer, e.g. follicular carcinoma.
 CC The PAX8-PPARGammael molecules are also useful for providing nucleotide
 CC and amino acid sequences useful for detecting the above disease. The
 CC present sequence represents a human PAX8e9-PPARGammael polypeptide.
 XX
 SQ Sequence 874 AA;

Query Match 60.0%; Score 2106.5; DB 22; Length 874;
 Best Local Similarity 83.7%; Pred. No. 2,7e-157;
 Matches 425; Conservative 4; Mismatches 14; Indels 65; Gaps 3;
 QY 1 MGETL-GDSPIDPES--DSFTDLSANISQEMTMVDTMEFMPFNIGISSVDSLVMEDS 57
 |||
 DB 367 VGETLGPYPPIPTSGQSVASAIAGVNAEMTMVDEMPFMTNIGISSVDSLVMEDS 426
 QY 58 HSDIPEFTTVDFSSISTPHIEDIPEPTRDPVYADKYDLKLOEYOSAIKVEPASPPYS 117
 |||
 DB 427 HSDIPEFTTVDFSSISTPHIEDIPEPTRDPVYADKYDLKLOEYOSAIKVEPASPPYS 486
 QY 118 EKTQLYN-----
 |||
 DB 487 EKTQLYNKKHHEPNSNLMAIECVGCDKASGFHYGVNACEGCKGFFRTIRLKLIDRCD 546
 QY 125 -----RNCQYCRFOKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDQLPESA 175
 |||
 DB 547 LNCRIHKKSRNKCQYCRFOKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDQLPESA 606
 QY 176 DLRALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFYIDMNSLMGEDKIKFKHITP 235
 |||
 DB 607 DLRALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFYIDMNSLMGEDKIKFKHITP 666
 QY 236 LQOSKEVAIRIFOGQCFRSEVAVOETETAKSIPGVNLDLNDQYTLTKYGVHEIITYM 295
 |||
 DB 667 LQOSKEVAIRIFOGQCFRSEVAVOETETAKSIPGVNLDLNDQYTLTKYGVHEIITYM 726
 QY 296 LASIMNKDGVLISEGQCFMTREFLSLRKPFQDMEPKFEFAVKFNALELDSDLAIFIA 355
 |||
 DB 727 LASIMNKDGVLISEGQCFMTREFLSLRKPFQDMEPKFEFAVKFNALELDSDLAIFIA 786
 QY 356 VIILSGDRPGLNVKPIEDIDQNLQALELQKLNPRESSQFAKLLQKMTDLRQIVTEH 415
 |||
 DB 787 VIILSGDRPGLNVKPIEDIDQNLQALELQKLNPRESSQFAKLLQKMTDLRQIVTEH 846
 QY 416 VQLLQVYIKKTEPDMSLPILQEIYKDL 443
 |||
 DB 847 VQLLQVYIKKTEPDMSLPILQEIYKDL 874
 RESULT 10
 AAB85793
 ID AAB85793 standard; Protein; 777 AA.
 XX
 AC AAB85793;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Human PAX8e7-PPARGammael protein sequence.
 XX
 KW PAX8-PPARGammael; oncogene; cytostatic; PAX8; PPARGammael; cancer;
 XX
 KW Follicular carcinoma; PAX8e7-PPARGammael; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200152789-A2.
 XX
 PD 26-JUL-2001.
 XX
 PF 18-JAN-2001; 2001WO-US01664.
 XX
 PR 20-JAN-2000; 2000US-0177109.
 XX
 PR 14-AUG-2000; 2000US-0225079.
 XX
 PA (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
 XX
 PI Kroll TG, Fletcher JA;
 XX
 DR WPI; 2001-514487/56.
 XX
 DR N-PSDB; AAH76281.
 XX
 PT New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and

PT treating certain tumors or cancers, e.g. follicular carcinoma -
XX
PS Claim 16; Page 100-104; 145pp; English.
XX
XX The invention relates to an oncogene designated PAX8-PPARgamma1 that
CC contains a PAX8 coding region fused to PPARgamma1 coding region. The
CC PAX8-PPARgamma1 polypeptides can be expressed by standard recombinant
CC methodology. A PPARgamma1 ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARgamma1, where the disorder is cancer, e.g. follicular carcinoma.
CC The PAX8-PPARgamma1 molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The
CC present sequence represents a human PAX8e7-PPARgamma1 polypeptide.
XX
SQ Sequence 777 AA:

Query Match 59.8%; Score 2099; DB 22; Length 777;
Best Local Similarity 83.3%; Pred. No. 8.8e-157;
Matches 420; Conservative 5; Mismatches 13; Indels 66; Gaps 2;

QY 6 GDSPIPESDSFTDTLSAN----ISOEMTVDTMPMPPTNGISSVDLSVMEDHSHSPD 61
DB 274 GKATLPSPNPLGRNLSTHOTYVVAEMTVDTMPMPPTNGISSVDLSVMEDHSHSPD 333
QY 62 IKPFTVDESSISTPHYEDIPFTRDPVADYKYDKLQGYQSAIKYEPASPPYSEKIQ 121
DB 334 IKPFTVDESSISTPHYEDIPFTRDPVADYKYDKLQGYQSAIKYEPASPPYSEKIQ 393
QY 122 LYN----- 124
DB 394 LYNKPHEEPSLSMAIECRVCGDKASGFHYGHVHACEGCKFFRTIRLKLIDRCDLNCR 453
QY 125 -----RNKCOYCRFOKCLAVGMSHNAIRFGMPQAEKEKILAEISSIDIDLPESADLRA 179
DB 454 IHKSNRKCQYCRFOKCLAVGMSHNAIRFGMPQAEKEKILAEISSIDIDLPESADLRA 513
QY 180 LAHKLIDSYTKSPPLTKAKARAILTGKTTDKSPFYIYDNMSLMKGEDKIKFKHITPLQEQ 239
DB 514 LAHKLIDSYTKSPPLTKAKARAILTGKTTDKSPFYIYDNMSLMKGEDKIKFKHITPLQEQ 573
QY 240 SKFVAIRIFGCGFRSVEAEOITEYAKSIPGFVNLDNQVTLKRGVHEIITYMLASL 299
DB 574 SKFVAIRIFGCGFRSVEAEOITEYAKSIPGFVNLDNQVTLKRGVHEIITYMLASL 633
QY 300 MNKGVYLISGCGFMTREFLSLRKPGDEMBKFEFAVKNALDELDDSLAIFIAVIL 359
DB 634 MNKGVYLISGCGFMTREFLSLRKPGDEMBKFEFAVKNALDELDDSLAIFIAVIL 693
QY 360 SGDRPGILANKPIEDTODNLQALQLKLNHPSSQLFAKLQKMTDLRQIYTEHVQL 419
DB 694 SGDRPGILANKPIEDTODNLQALQLKLNHPSSQLFAKLQKMTDLRQIYTEHVQL 753
QY 420 QVIKKTETDMSLHPLQEIYKDL 443
DB 754 QVIKKTETDMSLHPLQEIYKDL 777

RESULT 11
AAR9323
ID AAR9323 standard; protein; 494 AA.
XX
AC AAR9323;
XX
DT 14-APR-1997 (first entry)
XX
DE Peroxisome proliferator activated receptor gamma subtype 1.
XX
XX Peroxisome proliferator activated receptor gamma; PPARgamma; human;
KW peroxisome; cholesterol metabolism; lipid metabolism; respiration;
KW fatty acid metabolism; beta-oxidation cycle enzyme; hypolipidemic drug;
KW herbicide; hppargamma2; adipose tissue; obesity; diabetes; anorexia;
KW hyperlipidemia; hypercholesterolaemia; hyperlipoproteinaemia; antibody;
KW metabolic disorder; therapy.

XX
OS Homo sapiens.
XX
PN MO9623884-A2.
XX
PD 08-AUG-1996.
XX
PF 29-JAN-1996; 96WO-US01469.
XX
PR 23-OCT-1995; 95US-0005809.
PR 30-JAN-1995; 95US-0380051.
PR 07-JUN-1995; 95US-0484487.
XX
PA (LIGA-) LIGAND PHARM INC.
XX
PI Mukherjee R;
XX
DR WPI; 1996-384114/38.
DR N-PSDB; AAT35333.
XX
PT Human peroxisome proliferator activated receptor (s), hppar-gamma and
PT hppar-gamma2 - useful to identify polypeptide (ant)agonists to treat
e.g. obesity, anorexia and diabetes
XX
PS Claim 24; Page 85-87; 108pp; English.
XX
XX AAR9323-R9325 represent the three different subtypes of human
CC peroxisome proliferator activated receptor gamma (hPPARgamma). These
CC sequences differ in their N-terminal sequences. Peroxisomes contain
CC enzymes for cholesterol and lipid metabolism, and respiration.
CC Peroxisome proliferators increase the capacity of the peroxisomes to
CC metabolize fatty acids, via increased expression of the beta-oxidation
CC cycle enzymes. Peroxisome proliferators include unsaturated fatty
CC acids, hypolipidemic drugs, and herbicides. PPARgamma has two main
CC subtypes (hPPARgamma, and hPPARgamma2), which differ in their N-terminal
CC sequences, and are expressed at high levels in adipose tissue.
CC hPPARgamma proteins repress hpparalpha activity. These proteins can be
CC used in a cell system with a reporter gene (and optionally a hpparalpha
CC protein) to identify agonists and antagonists of the polypeptides, which
CC can be used as therapeutic agents. The agonists and antagonists can be
CC used to treat obesity, diabetes, anorexia, hyperlipidemia,
CC hypercholesterolaemia, hyperlipoproteinaemia, and other metabolic
CC disorders. These sequences, the DNA encoding them, and antibodies
CC against them, can be used to establish a tissue specific expression
CC pattern for hppargamma.
XX
SQ Sequence 494 AA:

Query Match 59.7%; Score 2094; DB 17; Length 494;
Best Local Similarity 87.0%; Pred. No. 1.2e-156;
Matches 416; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 28 EMTMVDTMPMPPTNGISSVDLSVMEDHSHSPDIPFTTVPDSSISTPHYEDIPFTRID 87
DB 17 EMTMVDTMPMPPTNGISSVDLSVMEDHSHSPDIPFTTVPDSSISTPHYEDIPFTRID 76
QY 88 PNVADYKYDKLQGYQSAIKYEPASPPYSEKIQ-----RNKCOYCRFOKCLAVGMSHNA 145
DB 77 PNVADYKYDKLQGYQSAIKYEPASPPYSEKIQ-----RNKCOYCRFOKCLAVGMSHNA 196
QY 125 -----RNKCOYCRFOKCLAVGMSHNA 145
DB 137 GFHYGVHACEGCKGFFRTIRLKLIDRCDLNCR.IHKSNRKCQYCRFOKCLAVGMSHNA 196
QY 146 IRFRMPQAEKEKILAEISSIDIDLPESADLRA.LAKHLIDSYTKSPPLTKAKARAILNG 205
DB 197 IRFRMPQAEKEKILAEISSIDIDLPESADLRA.LAKHLIDSYTKSPPLTKAKARAILNG 256
QY 206 KTDKSPFYIYDNMSLMKGEDKIKFKHITPLQEQSKFVAIRIFGCGFRSVEAEOITEY 265
DB 257 KTDKSPFYIYDNMSLMKGEDKIKFKHITPLQEQSKFVAIRIFGCGFRSVEAEOITEY 316
QY 266 AMSIPGFVNLDNQVTLKRGVHEIITYMLASLMNKDGVYLISGCGFMTREFLSLRK 325

```

Db 317 AKSIPGVNLDNDQVTLTKYGVHEIITYMLASLNKKGVLISEGQGMTRFEFLSKRP 376
OY 326 FGDFMEPKFEFVAVKFNALIEDSDLAIFIAVITILSGDRPGLLNKPRIDIDQNLQALEL 385
Db 377 FGDFMEPKFEFVAVKFNALIEDSDLAIFIAVITILSGDRPGLLNKPRIDIDQNLQALEL 436
OY 386 QLKLNHPSSQLFAKLQKMTDLROIVTEHVQLLOVIKKETDMSLHPLQETIKDLY 443
Db 437 QLKLNHPSSQLFAKLQKMTDLROIVTEHVQLLOVIKKETDMSLHPLQETIKDLY 494

RESULT 12
AAB85794
ID AAB85794 standard; Protein: 840 AA.
AC AAB85794;
XX
XX 29-OCT-2001 (first entry)
XX
XX Human PAX88-PPARGamma1 protein sequence.
XX
XX PAX8-PPARGamma1; oncogene; cytosolic; PAX8; PPARGamma1; cancer;
XX
XX follicular carcinoma; PAX88-PPARGamma1; human.
XX
XX Homo sapiens.
XX
XX WO200152789-A2.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US01664.
XX
XX 20-JAN-2000; 2000US-0177109.
XX
XX 14-AUG-2000; 2000US-0225079.
XX
XX (BGMH) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Krcoll TG, Fletcher JA:
XX
XX WPI: 2001-514487/56.
XX
XX N-PSDB: AAH76282.
XX
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
XX
XX treating certain tumors or cancers, e.g. follicular carcinoma
XX
XX Claim 16; Page 106-109; 145pp; English.
XX
XX The invention relates to an oncogene designated PAX8-PPARGamma1 that
XX
XX contains a PAX8 coding region fused to PPARGamma1 coding region. The
XX
XX PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant
XX
XX methodology. A PPARGamma1 ligand or agent is useful for treating a
XX
XX subject having a disorder characterized by the presence of a PAX8-
XX
XX PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.
XX
XX The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
XX
XX and amino acid sequences useful for detecting the above disease. The
XX
XX present sequence represents a human PAX88-PPARGamma1 polypeptide.
XX
XX Sequence 840 AA:
XX
XX Query Match 59.7%; Score 2094; DB 22; Length 840;
XX
XX Best Local Similarity 87.0%; Pred. No. 2.4e-156;
XX
XX Matches 416; Conservative 0; Mismatches 0; Indels 62; Gaps 1:
XX
OY 28 EKTMTVDTEPPTNFGISSVDLSVMEHSHSFDIKPTTYDFSSISPHHEDIPFTRTD 87
Db 363 EKTMTVDTEPPTNFGISSVDLSVMEHSHSFDIKPTTYDFSSISPHHEDIPFTRTD 422
OY 88 PVVADYKDKLQEQYSAIKVEPPASPPYSEKTOLYN----- 124
Db 423 PVVADYKDKLQEQYSAIKVEPPASPPYSEKTOLYNKPHEPPSSLSMAIECRVCGDKAS 482
OY 125 -----RNKQYCRFOKCLAVGMSHNA 145

```

```

Db 483 GFHYGVHACEGCKGFRRTRIRKLTIYDRCDLNCRIHKKSRNKCQCRFOKCLAVGMSHNA 542
OY 146 IRFGMPQAEKEKLAIESSDIDQNPESADIRALAKLYDSYISFPLTAKARAILTG 205
Db 543 IRFGMPQAEKEKLAIESSDIDQNPESADIRALAKLYDSYISFPLTAKARAILTG 602
OY 206 KTTDPSPIVIDMNSLMGEDIKFKHITPLOEQSEKVAIRFPGQCRPSVAOETIEY 265
Db 603 KTTDPSPIVIDMNSLMGEDIKFKHITPLOEQSEKVAIRFPGQCRPSVAOETIEY 662
OY 266 AKSIPGVNLDNDQVTLTKYGVHEIITYMLASLNKKGVLISEGQGMTRFEFLSKRP 325
Db 663 AKSIPGVNLDNDQVTLTKYGVHEIITYMLASLNKKGVLISEGQGMTRFEFLSKRP 722
OY 326 FGDFMEPKFEFVAVKFNALIEDSDLAIFIAVITILSGDRPGLLNKPRIDIDQNLQALEL 385
Db 723 FGDFMEPKFEFVAVKFNALIEDSDLAIFIAVITILSGDRPGLLNKPRIDIDQNLQALEL 782
OY 386 QLKLNHPSSQLFAKLQKMTDLROIVTEHVQLLOVIKKETDMSLHPLQETIKDLY 443
Db 783 QLKLNHPSSQLFAKLQKMTDLROIVTEHVQLLOVIKKETDMSLHPLQETIKDLY 840

RESULT 13
AAR99324
ID AAR99324 standard; Protein: 477 AA.
AC AAR99324;
XX
XX 14-APR-1997 (first entry)
XX
XX Peroxisome proliferator activated receptor gamma subtype 2.
XX
XX Peroxisome proliferator activated receptor gamma; PPARGamma; human;
XX
XX peroxisome; cholesterol metabolism; lipid metabolism; respiration;
XX
XX fatty acid metabolism; beta-oxidation cycle enzyme; hypolipidemic drug;
XX
XX herbicide; hPPARGamma2; adipose tissue; obesity; diabetes; anorexia;
XX
XX hyperlipidaemia; hypercholesterolaemia; hyperlipoproteinaemia; antibody;
XX
XX metabolic disorder; therapy.
XX
XX Homo sapiens.
XX
XX WO9623884-A2.
XX
XX 08-AUG-1996.
XX
XX 29-JAN-1996; 96WO-US01469.
XX
XX 23-OCT-1995; 95US-0005809.
XX
XX 30-JAN-1995; 95US-0380051.
XX
XX 07-JUN-1995; 95US-0484487.
XX
XX (LIGA-) LIGAND PHARM INC.
XX
XX Mukherjee R:
XX
XX WPI: 1996-384114/38.
XX
XX N-PSDB: AAT35333.
XX
XX Human peroxisome proliferator activated receptor(s), hPPAR-gamma and
XX
XX hPPAR-gamma2 - useful to identify polypeptide (ant)agonists to treat
XX
XX e.g. obesity, anorexia and diabetes
XX
XX Claim 25; : 108pp; English.
XX
XX AAR99323-R99325 represent the three different subtypes of human
XX
XX peroxisome proliferator activated receptor gamma (hPPARGamma). These
XX
XX sequences differ in their N-terminal sequences. Peroxisomes contain
XX
XX enzymes for cholesterol and lipid metabolism, and respiration.
XX
XX Peroxisome proliferators increase the capacity of the peroxisomes to
XX
XX metabolise fatty acids, via increased expression of the beta-oxidation
XX
XX cycle enzymes. Peroxisome proliferators include unsaturated fatty

```

CC acids, hypolipidemic drugs, and herbicides. PPARgamma has two main
 CC subtypes (hPPARgamma, and hPPARgamma2), which differ in their N-terminal
 CC sequences, and are expressed at high levels in adipose tissue.
 CC hPPARgamma proteins repress hPPARalpha activity. These proteins can be
 CC used in a cell system with a reporter gene (and optionally a hPPARalpha
 CC protein) to identify agonists and antagonists of the polypeptides, which
 CC can be used as therapeutic agents. The agonists and antagonists can be
 CC used to treat obesity, diabetes, anorexia, hyperlipidaemia,
 CC hypercholesterolaemia, hyperlipoproteinaemia, and other metabolic
 CC disorders. These sequences, the DNA encoding them, and antibodies
 CC against them, can be used to establish a tissue specific expression
 CC pattern for hPPARgamma.

CC
 XX
 SQ Sequence 477 AA:

Query Match 59.5% Score 2089; DB 17; Length 477;

Best Local Similarity 87.0% Pred. No. 2.8e-156; Mismatches 0; Indels 62; Gaps 1;

Matches 415; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 29 MTMVDTEMPFPTNFGISSVDLSVMDHSHSPDIKPFYTVDFSSISTPHYEDIPTRTDP 88

DB 1 MTMVDTEMPFPTNFGISSVDLSVMDHSHSPDIKPFYTVDFSSISTPHYEDIPTRTDP 60

QY 89 VVADKYDLDKLOEYOSAIKVEPASPYYSEKTQLYN----- 124

DB 61 VVADKYDLDKLOEYOSAIKVEPASPYYSEKTQLYNKPHEPSNSLMAIECRVCGDKASG 120

QY 125 -----RNCQYCRFQKCLAVGSHNAI 146

DB 121 FHGYHACGCGGFFRRTIRLKIYDRCLNCRIRKSKNNKQYCRFQKCLAVGSHNAI 180

QY 147 RGRMPQAEKEKLLAEISSDIDQLNPEASADLRALKHLYDSYIKSFPLTKAKARAILTGK 206

DB 181 RGRMPQAEKEKLLAEISSDIDQLNPEASADLRALKHLYDSYIKSFPLTKAKARAILTGK 240

QY 207 TTDKSPFYIYDNNLSLMGMDKIKFKHITPLOSKEVAIRIQGCGFRSVEAVOETTEYA 266

DB 241 TTDKSPFYIYDNNLSLMGMDKIKFKHITPLOSKEVAIRIQGCGFRSVEAVOETTEYA 300

QY 267 KSIPGFVNLDNDQVTLKYGVEHIIYTMASLMNKDGLVISEGCGFMTREFLSLRKPF 326

DB 301 KSIPGFVNLDNDQVTLKYGVEHIIYTMASLMNKDGLVISEGCGFMTREFLSLRKPF 360

QY 327 GDFMEPKFEFAVKFNALDSDLAIFIAVILISGDRGLNVKPIEDIQDNLQALELQ 386

DB 361 GDFMEPKFEFAVKFNALDSDLAIFIAVILISGDRGLNVKPIEDIQDNLQALELQ 420

QY 387 LKLNHPESQLEAKLLQKTDLRQYITENVOLLOVYIKKETMSLHPLLQETIKDLY 443

DB 421 LKLNHPESQLEAKLLQKTDLRQYITENVOLLOVYIKKETMSLHPLLQETIKDLY 477

RESULT 14

AAR9327 standard: Protein; 477 AA.

AC AAR9327;

DT 14-Apr-1997 (first entry)

XX Peroxisome proliferator activated receptor gamma2 subtype 2.

XX Peroxisome proliferator activated receptor gamma; PPARgamma; human;

KM peroxisome; cholesterol metabolism; lipid metabolism; respiration;

KM fatty acid metabolism; beta-oxidation cycle enzyme; hypolipidemic drug;

KM herbicide; hPPARgamma2; adipose tissue; obesity; diabetes; anorexia;

KM hyperlipidaemia; hypercholesterolaemia; hyperlipoproteinaemia; antibody;

XX metabolic disorder; therapy.

XX Homo sapiens.

XX OS

XX PN MO9623884-A2.

PD 08-AUG-1996.

PF 29-JAN-1996; 96MO-US01469.

PR 23-OCT-1995; 95US-0005809.

PR 30-JAN-1995; 95US-0380051.

PR 07-JUN-1995; 95US-0484487.

PA (LIGA-) LIGAND PHARM INC.

PI Mukherjee R;

DR WPI; 1996-384114/38.

DR N-PSDB; AAT35334.

PT Human peroxisome proliferator activated receptor(s), hPPAR-gamma and

PT hPPAR-gamma2 - useful to identify polypeptide (ant)agonists to treat

PT e.g. obesity, anorexia and diabetes

PS Claim 43; : 108pp; English.

CC AAR9326-R99328 represent the three different subtypes of human

CC peroxisome proliferator activated receptor gamma2 (hPPARgamma2). These

CC sequences differ in their N-terminal sequences. Peroxisomes contain

CC enzymes for cholesterol and lipid metabolism, and respiration.

CC Peroxisome proliferators increase the capacity of the peroxisomes to

CC metabolise fatty acids, via increased expression of the beta-oxidation

CC cycle enzymes. Peroxisome proliferators include unsaturated fatty

CC acids, hypolipidemic drugs, and herbicides. PPARgamma has two main

CC subtypes (hPPARgamma, and hPPARgamma2), which differ in their N-terminal

CC sequences, and are expressed at high levels in adipose tissue.

CC hPPARgamma proteins repress hPPARalpha activity. These proteins can be

CC used in a cell system with a reporter gene (and optionally a hPPARalpha

CC protein) to identify agonists and antagonists of the polypeptides, which

CC can be used as therapeutic agents. The agonists and antagonists can be

CC used to treat obesity, diabetes, anorexia, hyperlipidaemia,

CC hypercholesterolaemia, hyperlipoproteinaemia, and other metabolic

CC disorders. These sequences, the DNA encoding them, and antibodies

CC against them, can be used to establish a tissue specific expression

CC pattern for hPPARgamma.

SQ Sequence 477 AA:

Query Match 59.5% Score 2089; DB 17; Length 477;

Best Local Similarity 87.0% Pred. No. 2.8e-156; Mismatches 0; Indels 62; Gaps 1;

Matches 415; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 29 MTMVDTEMPFPTNFGISSVDLSVMDHSHSPDIKPFYTVDFSSISTPHYEDIPTRTDP 88

DB 1 MTMVDTEMPFPTNFGISSVDLSVMDHSHSPDIKPFYTVDFSSISTPHYEDIPTRTDP 60

QY 89 VVADKYDLDKLOEYOSAIKVEPASPYYSEKTQLYN----- 124

DB 61 VVADKYDLDKLOEYOSAIKVEPASPYYSEKTQLYNKPHEPSNSLMAIECRVCGDKASG 120

QY 125 -----RNCQYCRFQKCLAVGSHNAI 146

DB 121 FHGYHACGCGGFFRRTIRLKIYDRCLNCRIRKSKNNKQYCRFQKCLAVGSHNAI 180

QY 147 RGRMPQAEKEKLLAEISSDIDQLNPEASADLRALKHLYDSYIKSFPLTKAKARAILTGK 206

DB 181 RGRMPQAEKEKLLAEISSDIDQLNPEASADLRALKHLYDSYIKSFPLTKAKARAILTGK 240

QY 207 TTDKSPFYIYDNNLSLMGMDKIKFKHITPLOSKEVAIRIQGCGFRSVEAVOETTEYA 266

DB 241 TTDKSPFYIYDNNLSLMGMDKIKFKHITPLOSKEVAIRIQGCGFRSVEAVOETTEYA 300

QY 267 KSIPGFVNLDNDQVTLKYGVEHIIYTMASLMNKDGLVISEGCGFMTREFLSLRKPF 326

DB 301 KSIPGFVNLDNDQVTLKYGVEHIIYTMASLMNKDGLVISEGCGFMTREFLSLRKPF 360

QY 327 GDFMEPKFEFAVKFNALDSDLAIFIAVILISGDRGLNVKPIEDIQDNLQALELQ 386

Db 361 GDFMEPKFEFAVKFNALELDSDLAIFIAVITLLSGDRGLNVKPIEDIQDNLQALELQ 420
QY 387 LKLNHPSSQLFAKLLQKMTDLROIVTEHVOLLQYIKKTEFDMSLHPLQETIKDLY 443
|||||
Db 421 LKLNHPSSQLFAKLLQKMTDLROIVTEHVOLLQYIKKTEFDMSLHPLQETIKDLY 477
RESULT 15
AAE12867
ID AAE12867 standard; Protein: 477 AA.
XX
AC AAE12867;
XX
DT 15-JAN-2002 (first entry)
XX
DE Human peroxisome proliferator activated receptor (PPAR) gamma1 protein.
XX
KM Human; peroxisome proliferator activated receptor gamma1; PPAR gamma1;
KM cytoxic; antiproliferative; antiviral; cancer; cell proliferation;
KM viral infection; pharmaceutical; thiazolidinedione.
OS Homo sapiens.
XX
PN US6294559-B1.
XX
PD 25-SEP-2001.
XX
PF 03-AUG-1998; 98US-0128142.
XX
PR 02-MAY-1996; 96US-016694P.
PR 18-APR-1997; 97US-0844007.
XX
PA (MERI) MERCK & CO INC.
XX
PI Smith RG;
XX
DR WPI: 2001-647265/74.
DR N-PSDB; AAD21021.
XX
PT Use of thiazolidinedione for treating cancer and viral infections -
XX
PS Example 3; Fig 2; 17pp; English.
XX
CC The invention relates to compounds and ligands that bind to human
CC peroxisome proliferator activated receptors (PPAR) gamma1 and gamma2.
CC The invention is useful for treating cancer and other disorders including
CC excessive cell proliferation and viral infection. The invention is also
CC directed to the use of PPAR gamma1 and gamma2 to identify compounds that
CC are antiproliferative, antiviral and antitumor agents. The invention
CC also relates to a method of treating cancer using a pharmaceutical
CC composition comprising thiazolidinedione in an amount sufficient to
CC modulate PPAR gamma1 and gamma2 activity. The present sequence is human
CC peroxisome proliferator activated receptor (PPAR) gamma1 protein related
CC to the invention.
XX
SQ Sequence 477 AA;
Query Match 59.5%; Score 2089; DB 22; Length 477;
Best Local Similarity 87.0%; Pred. No. 2.8e-156;
Matches 415; Conservative 0; Mismatches 0; Indels 62; Gaps 1;
QY 29 MTMVDTEPFMPFTNGISSVDLSVMEHSHSFDIKPFTTVDFSSISSTPHYEDIPFTRDP 88
|||||
Db 1 MTMVDTEPFMPFTNGISSVDLSVMEHSHSFDIKPFTTVDFSSISSTPHYEDIPFTRDP 60
QY 89 VVADKYDLKLOEYOSAIKVEPASPPIYSEKTQLYN----- 124
|||||
Db 61 VVADKYDLKLOEYOSAIKVEPASPPIYSEKTQLYNKPHPEPSNSLMAIECRVCGDKASG 120
QY 125 -----RNKQYCRFQKCLAVGSHNAI 146
|||||
Db 121 FHVGHAGECGKFFRRITRLKLIYDRCDLNCRIHKSRNKCQYCRFQKCLAVGSHNAI 180

QY 147 RFRMPQAEKEKLLAEISSDIDQLNPESSADLRALAKHLVDSYIKSFPLYTKAKARAILTGK 206
|||||
Db 181 RFRMPQAEKEKLLAEISSDIDQLNPESSADLRALAKHLVDSYIKSFPLYTKAKARAILTGK 240
QY 207 TTDSKSPFIYDMNSLMGEEKIKRKHITPLOBOSKEVAIRIFOGCQFRSVAVOETIEYA 266
|||||
Db 241 TTDSKSPFIYDMNSLMGEEKIKRKHITPLOBOSKEVAIRIFOGCQFRSVAVOETIEYA 300
QY 267 KSIPGFVNLNDQVYTLKYGVHEIITYTMLASLMNKDGLTSEGOGEFMTREFLSLRKF 326
|||||
Db 301 KSIPGFVNLNDQVYTLKYGVHEIITYTMLASLMNKDGLTSEGOGEFMTREFLSLRKF 360
QY 327 GDFMEPKFEFAVKFNALELDSDLAIFIAVITLLSGDRGLNVKPIEDIQDNLQALELQ 386
|||||
Db 361 GDFMEPKFEFAVKFNALELDSDLAIFIAVITLLSGDRGLNVKPIEDIQDNLQALELQ 420
QY 387 LKLNHPSSQLFAKLLQKMTDLROIVTEHVOLLQYIKKTEFDMSLHPLQETIKDLY 443
|||||
Db 421 LKLNHPSSQLFAKLLQKMTDLROIVTEHVOLLQYIKKTEFDMSLHPLQETIKDLY 477

Search completed: February 25, 2003, 03:58:26
Job time : 65 secs

THIS PAGE BLANK (USPT)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 25, 2003, 03:59:06 ; Search time 368 Seconds
(without alignments)
4210.258 Million cell updates/sec

Title: US-09-931-007A-1
Perfect score: 3508
Sequence: 1 MGRTIGSDPIDPESDSFTDT.....KTEHDSIMPLIDYKLDY 688

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPRO.spool/US09931007/rnat_14022003_102553_4668/app_query.fasta_1.839
-DB=N_Geneseq_101002 -QFRT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsu62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09931007 @CGN.1.1.236 @rnat_14022003_102553_4668 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAR -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARP.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: N_Geneseq_101002:*
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2231	63.6	1518	22	AAD21022	Human peroxisome p
2	2231	63.6	1608	22	AAH76288	Human PPARgamma cd
3	2231	63.6	1608	22	AAF23644	Human PPARgamma co
4	2231	63.6	1647	17	AAT35334	Peroxisome prolif
5	2106.5	60.0	2596	22	AAH76294	Human PAX8b(- exon
6	2106.5	60.0	2625	22	AAH76283	Human PAX8b(-PPARg
7	2099	59.8	2334	22	AAH76281	Human PAX8b7-PPARg
8	2098	59.8	1936	17	AAT35333	Peroxisome prolif
9	2094	59.7	2523	22	AAH76282	Human PAX8b-PPARg
10	2089	59.5	1434	20	AAD21021	Human peroxisome p
11	2065.5	58.9	1811	20	AAK36522	Human PPAR gamma1
12	2065.5	58.9	1811	22	AAH76296	Human PPARgamma cd
13	2065.5	58.9	1811	24	ABN95714	Gene #2212 used to
14	2065.5	58.9	1844	22	AAF55663	Human peroxisome p
15	2065.5	58.9	2295	22	AA544653	Human full-length
16	2065	58.9	2005	17	AAK25481	Human peroxisome p
17	2065	58.9	2005	17	AAH76205	Peroxisome prolif
18	2065	58.9	2005	20	AAK29814	Mouse peroxisome p
19	1570	44.8	1278	22	AAH33499	Human colon cancer
20	1101.5	31.4	1326	17	AAT10583	Peroxisome prolif
21	1101.5	31.4	3294	14	AA048939	Steroid hormone re
22	1101.5	31.4	3301	20	AAK36523	Human PPAR-delta c
23	1101.5	30.8	1850	22	ABL65244	Lung cancer relate
24	1079.5	30.8	1407	16	AAQ30517	Peroxisome prolif
25	1076.5	30.7	2012	17	AAQ92081	Human peroxisome p
26	1076.5	30.7	2012	20	AAK29815	Peroxisome prolif
27	1076.5	30.7	1854	20	AAK36521	Mouse peroxisome p
28	1075.5	30.7	2009	14	AAQ39089	Human PPAR-alpha c
29	1056	30.1	732	20	AAK19070	XR4 coding sequenc
30	761	21.7	498	24	ABO59148	Human PPAR gamma 1
31	618.5	17.6	415	21	AAC98246	Human colon cancer
32	562	16.0	373	19	AAV21207	Human colon cancer
33	521	14.9	315	19	AAV21207	Mouse PPAR-gamma A
34	512.5	14.6	611	21	AAC98099	Human colon cancer
35	486	13.9	589	24	ABO60211	Human colon cancer
36	472	13.5	451	22	ABA48744	Human breast cell
37	472	13.5	451	22	ABA66654	Human foetal liver
38	472	13.5	451	22	ABA33719	Probe #12185 for g
39	472	13.5	451	22	AAK15083	Human brain expres
40	472	13.5	451	22	AAK40811	Human bone marrow
41	472	13.5	451	22	AAI21581	Probe #11514 for g
42	472	13.5	451	22	AAI6870	Probe #15556 used
43	472	13.5	451	22	AAI07275	Probe #7266 used t
44	472	13.5	451	24	ABSI4765	Human genome-deriv
45	435	12.4	576	24	ABO57991	Human colon cancer

ALIGNMENTS

RESULT 1
AAD21022
AAD21022 standard; CDNA: 1518 BP.
AAD21022;
15-JAN-2002 (first entry)
Human peroxisome proliferator activated receptor (PPAR) gamma2 CDNA.
Human; peroxisome proliferator activated receptor gamma2; PPAR gamma2;
cytotoxic; antiproliferative; antiviral; cancer; cell proliferation;
viral infection; pharmaceutical; thiazolidinedione; ss.
Homo sapiens.
Key Location/Qualifiers
CDS 1..1518

```
FT      /*tag= a
FT      /product= "Human peroxisome proliferator activated
FT      receptor (PPAR) gamma2 protein"
PN      US6294559-B1.
XX      25-SEP-2001.
XX      PD
XX      03-AUG-1998; 98US-0128142.
XX      PF
XX      02-MAY-1996; 96US-016694P.
XX      PR 18-APR-1997; 97US-0844007.
XX      PA (MERI ) MERCK & CO INC.
XX      PI Smith RG;
XX      DR WPI: 2001-647265/74.
XX      DR P-PSDB: AAE12868.
XX      PS
XX      Use of thiazolidinedione for treating cancer and viral infections -
XX      Example 3; Fig 3; 17pp; English.
XX      The invention relates to compounds and ligands that bind to human
XX      CC peroxisome proliferator activated receptors (PPAR) gamma1 and gamma2.
XX      CC The invention is useful for treating cancer and other disorders including
XX      CC excessive cell proliferation and viral infection. The invention is also
XX      CC directed to the use of PPAR gamma1 and gamma2 to identify compounds that
XX      CC are antiproliferative, antiviral and antitumour agents. The invention
XX      CC also relates to a method of treating cancer using a pharmaceutical
XX      CC composition comprising thiazolidinedione in an amount sufficient to
XX      CC modulate PPAR gamma1 and gamma2 activity. The present cDNA sequence
XX      CC encodes human peroxisome proliferator activated receptor (PPAR) gamma2
XX      CC protein related to the invention.
XX      SQ Sequence 1518 BP; 438 A; 363 C; 346 G; 371 T; 0 other:

Alignment Scores:
Pred. No.: 1.74e-197 Length: 1518
Score: 2231.00 Matches: 443
Percent Similarity: 87.72% Conservative: 0
Best Local Similarity: 87.72% Mismatches: 0
Query Match: 63.60% Indels: 62
DB: Gaps: 1

US-09-931-007A-1 (1-688) x AMD21022 (1-1518)
QY 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20
DB 1 ATGGGTGAAACTCTGGAGATTCTCTATTGGACCCAGAAAGCGATTCTTCCTACTGATACA 60
QY 21 LeuSerAlaAsnIleSerGlnIleMetThrMetValAspThrGluMetProPheThrPro 40
DB 61 CTGTGTGCAACATATTCACAAAGAAATGACCATTGTTGACACAGAGATGCGCATTTGGCCC 120
QY 41 ThrAsnPhenGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
DB 121 ACCAACTTTGGGATCAGCTCCGTCGATCTCTCCGTAATGAGAACACCCACCTCCTTT 180
QY 61 AspIleLeuSerProPheThrValAspPheSerSerIleSerThrProHisThrGluAsp 80
DB 181 GATATCAAGCCCTTCCTACTGTTGACTTCTCCACATTTCCTACCTCCATPACGAAGAC 240
QY 81 IleProPheThrArgThrAspProValAlaAspTyrIleTyrAspLeuIleGln 100
DB 241 ATTCATTTCACAGAACACGATCCAGTGTGTGAGATTCAAGATATGACCTGAAACTTCA 300
QY 101 GluTyrGlnSerAlaIleIleValGluProAlaSerProTyrTyrSerGluIleThr 120
DB 301 GAGTACCAAGAGCAATCAATGAGAGCTGATCTCCACCTTATTATTTCAGAAAGACT 360
QY 121 GlnLeuTyrAsn----- 124
```

```
DB 361 CAGCTCTAACATTAAGCTTCATGAAGAGCTTCCAACTCCCTCATGGCAATTGATGTCT 420
QY 124 ----- 124
DB 421 GTCGTGGAGATTAAGCTTCTCGATTTCACATAGAGATTCAATGCTTGTGAAGATGCAAG 480
QY 124 ----- 124
DB 481 GGTTCCTCCGGAAGAACATCAATGAAAGCTTATCTATGACAGATGATCTTAAGTGT 540
QY 125 -----ArgAsnIleScyGlnTyrCysArgPheGlnIleScyLeuAla 138
DB 541 CCGATCCACAAAAAAGTGAATTAATGTCACTGCTCGGTTTCAGAAATAGCTTCA 600
QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluIleScyLeu 158
DB 601 GTGGGATGCTCATTAATGCCATGAGTTTGGCGGATGCCACAGCCGAGAAAGAGAG 660
QY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
DB 661 CTGTTGGCGGAGATCTCCAGTATATCCAGCCAGCTGAATCCAGAGTCCGCTGACCTCCGG 720
QY 179 AlaLeuAlaIleHisLeuIleAspSerTyrIleIleYSerPheProLeuThrLysAlaIle 198
DB 721 GCCCTGGCAAAACATTTGATGATCATATCAATTAAGTCTTCCCTGACCAAGCAAG 780
QY 199 AlaArgAlaIleLeuThrGlyIleThrArgPheIleAspIleThrPheValIleTyrAspMet 218
DB 781 GCGAGGGCGATTTTCACAGGAAGAACACAGCAATACCACTTGTATATATGACATG 840
QY 219 AsnSerLeuMetLeuGlyIleAspIleIleYSerPheIleThrProLeuGlnIle 238
DB 841 AATTCCTTAATGATGGAGAAATCAATCAAGTTCAACACATGACCCCTCGAGAG 900
QY 239 GlnSerIleGluValAlaIleArgIlePheGlnIleCysGlnPheArgSerValAlaIle 258
DB 901 CAGACCAAGAGAGTGGCCATCCGATCTTTCAGGCTCCCAATTCGCTCCGTGAGAGCT 960
QY 259 ValGlnGluIleThrGluTyrAlaIleYSerIleProGluPheValAsnLeuAspLeuAsn 278
DB 961 GTCCAGAGATACAGAGATGATGCCAAAAGCATCTCTGTTTGTAAATCTTGACTTGCAAC 1020
QY 279 AspGlnValThrLeuLeuIleYSerIleGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
DB 1021 GACCAAGTAATCTCTCAATATGAGAGTCCAGAGATCATTTACACATGCTGCTGCC 1080
QY 299 LeuMetAsnIleAspGlyValLeuIleSerGlnIleGlyGlnIlePheMetThrArgGluPhe 318
DB 1081 TTGATGAATTAAGAGATGGGTTCTCATATCCGAGGCCAAGGCTTCATGACAAAGGAGTTT 1140
QY 319 LeuIleYSerLeuArgIleProPheGlyAspPheMetGluProIlePheGluPheAlaIle 338
DB 1141 CTTAAAGAGCTCTCGAAGGCTTTTGTGACTTTATGAGAGCCCAAGTTGAGTTGCTGTG 1200
QY 339 LysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle 358
DB 1201 AAGTTCATATGCACTGTGATTTAGATGACAGCACTTGGCAATATTTATTTGCTGCTATT 1260
QY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValIleYSerProIleGluAspIleGlnAspAsn 378
DB 1261 CTCAGTGGAGACCGCCGAGGTTTGTGTAATGGAAGCCCATTTGAAGACATTCAAAGCAAC 1320
QY 379 LeuLeuGlnAlaLeuGlnLeuGlnLeuIleYSerAsnHisProGluYSerSerGlnLeuPhe 398
DB 1321 CTGCTACAAAGCCCTGGAGCTCCAGCTGAAGCTGAAGCCACCTGAGTCTCCACAGCTGTT 1380
QY 399 AlaIleLeuLeuGlnIleMetThrAspLeuArgGlnIleValThrGluHisValGlnLeu 418
DB 1381 GCCAAGCTGCTCCAAATAATGACAGACTCAGACAGATTTGACAGGAGACAGTGGAGCTA 1440
QY 419 LeuGlnValIleIleYSerIleGluThrAspMetSerLeuHisProLeuLeuGlnIle 438
```


Dd	1441	CTGCAGGCGATCAGAAACGACGACAGCATGACTTTCACCCGGCTCTCGAGAGATC	1500
Oy	439	TyrTyAspLeuTyr 443 	
Dd	1501	TACACGAGACTGTAC	1515
RESULT 2			
ID	AAH76288		
XX	AAH76288 standard; cDNA; 1608 BP.		
AC	AAH76288;		
DT	29-OCT-2001 (first entry)		
DE	Human PPARgamma cDNA sequence.		
XX			
KW	PAX8-PPARGamml; oncogene; cytotstatic; PAX8; PPARgamml; cancer; follicular carcinoma; PPARgamma; human; ss.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	91..1608 /*tag= a	
PN	WO200152789-A2.		
PD	26-JUL-2001.		
XX	18-JAN-2001; 2001WO-USO1664.		
PR	20-JAN-2000; 2000US-0177109.		
PR	14-AUG-2000; 2000US-0225079.		
PA	(BGHM) BRIGHAM & WOMENS HOSPITAL INC.		
PI	Kroll TG, Fletcher JA;		
DR	WPI; 2001-514487/56.		
DR	P-PSDB; AAB85800.		
PT	New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and treating certain tumors or cancers, e.g. follicular carcinoma -		
PS	Disclosure; Page 123-125; 145pp; English.		
XX			
CC	The invention relates to an oncogene designated PAX8-PPARGamml that contains a PAX8 coding region fused to PPARgamml coding region. The PAX8-PPARGamml polypeptides can be expressed by standard recombinant methodology. A PPARgamma ligand or agent is useful for treating a subject having a disorder characterized by the presence of a PAX8- PPARGamml, where the disorder is cancer, e.g. follicular carcinoma. The PAX8-PPARGamml molecules are also useful for providing nucleotide and amino acid sequences useful for detecting the above disease. The present sequence represents a human PPARgamma polypeptide encoding cDNA		
SQ	Sequence 1608 BP; 462 A; 380 C; 361 G; 405 T; 0 other;		
Alignment Scores:			
Pred. No.:	1.87e-197	Length:	1608
Score:	2231.00	Matches:	443
Percent Similarity:	87.72%	Conservative:	0
Best Local Similarity:	87.72%	Mismatches:	0
Query Match:	63.60%	Indels:	62
DB:	22	Gaps:	1
US-09-931-007A-1 (1-688) x AAH76288 (1-1608)			
Oy	1	MetGLYIuThrLeuGLYAspSerProIIleasProgiUserASpSerPhethrAspThr 20 	
Dd	91	ATGGTGAAACTCTGGGAGATTCTTCATTACACCAGAAACGATTCTTACTACGATAACA 150 	
Oy	21	LeuSerAlaAsnIleSerGlnGlumetThrMetValAsprhrglunmetPropheTrpPro 40 	

Db	151	CTGTCGCAAAACATATCCACGAAGAATGACCAATGGTTGACACAGAGATGCCATTCCTGGCCC	210
Qy	41	ThrsApnPhcglyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe	60
Db	211	ACCAACTTTGGGATGAGCTCCGCTGGATCTCTCCGTAATGGAAAGCCACTCCCACTCCTTT	270
Qy	61	AspIleLysProPheThrThrValAspPheSerIleSerThrProHisTryGluAsp	80
Db	271	GATATCAAGACCCTTCACTACTGTGTGACTTCTCCAGCAATTTCTACTCCACATTCAGAAAGAC	330
Qy	81	IleProPheThrArgThrAspProValValAlaAspIlyIlyTryAspLeuLysLeuGln	100
Db	331	ATTCCATTCCACAGAAACAGATCCAGGTGGTTGCACATTTACAAAGTATGACCGAAACCTTCAA	390
Qy	101	GluTryGlnSerAlaIleLysValGlnProIleAspPropotyrrTrySerGluLysThr	120
Db	391	GAGTCCAAAGTGCATCATTAAGTGGAGGCTTCATCTCCACCTTATTATTTCTGAGAAAGACT	450
Qy	121	GlnLeuTryAsn-----	124
Db	451	CAGCTCTTCATTAAGCCCTCATGAAAGAGCTTCCAAATCCCATGGCAATTCGATATGATGCT	510
Qy	124	-----	124
Db	511	GCTGTGGAGATTAAGCTTCTGCATTTCCATGAGACTTCATGCTTGTGAAGATGCAAG	570
Qy	124	-----	124
Db	571	GGTTTCTTCCGGAGAACATCAGATTGAGCCGTATCTATGACAGATGTGATCTTAAGCT	630
Qy	125	-----ArgAsnLysCysGlnTryCysArgPheGlnLysCysLeuAla	138
Db	631	CGGATCCACAAAAGAAAGTAGAAATTAATGATGACTGCTGGTTTCCGAAGATGCCCTTGCA	690
Qy	139	ValGlyMetSerHisAsnAlaIleArgPheGlyIleArgMetProGlnAlaGluLysGluLys	158
Db	691	GTGGGAGTGTCTCAATAATGCATCAGGTTTGGCGGATGCCACAGCGCGAAGAGAGAG	750
Qy	159	LeuLeuAlaGluIleLeuSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg	178
Db	751	CTGTTGGCGGAGATCTCCAGAGATATCCAGCAGCTGATCCAGAGTCCGCTGACCTTCGCG	810
Qy	179	AlaLeuAlaLysHisLeuTryAspSerTryIleLysSerPheProLeuThrValAlaLys	198
Db	811	GGCCTGGCAAAACATTTTATGACTCATACATAAAGTCTTCCGCTGACCAAGCAAG	870
Qy	199	AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTryAspMet	218
Db	871	GGAGGGGAGATCTTGACAGAAAGCAACACAAATCCACATTCGTTATCTATGACATG	930
Qy	219	AsnSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGlu	238
Db	931	ATTTCCTTAATGATGGGGAAGATTAATTAAGTTCAAGTTCAAACATCACACCCCTGCGAGAG	990
Qy	239	GlnSerLysGluValAlaIleArgIlePheGlnLysCysGlnPheArgSerValGluAla	258
Db	991	CAGACCAAAAGAGTGGCCATCCGATCTTTCAGGGCTGCCAGTTTCCCTCCGTGGAGCT	105
Qy	259	ValGlnGluIleThrGluTryAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn	278
Db	1051	GTGCGAGGAGTACACAGAGTATGCCAAAGCATCTCTGCTTTTGAATCTTGACTTGAAAC	111
Qy	279	AspGlnValThrLeuLeuLysTryGlyValHisGlnIleIleIleTryThrMetLeuAlaSer	298
Db	1111	GACCAAGTAACTCTCTCAATATGAGGTCCACAGATCATTTACAAATAGCTGGCCCTCC	117
Qy	289	LeuMetAsnLysAspGlyValLeuIleSerGlnLysGlnGlyPheMetThrArgGluPhe	318
Db	1171	TTGATGATTAAGATGGGGTTCTCATATCCAGGGGCCAAGGCTTCATGACAAAGGACTTT	123
Qy	319	LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal	338

Db 1231 CTAAGACCTCGGAAGCCTTTGGTGACTTTATGAGACCCCAAGTTTGACTTTCCTGTG 1290
QY 339 LysPheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIleIle 358
Db 1291 AAGTTCATGCGACTGGAAATGATGACAGCGACTGGCAATTTTATTTGCTCATTTATT 1350
QY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
Db 1351 CTCAGTGGAGACCGCCGACGTTTGGTGAATGTGAAGCCATTGAAGACATTCAGACAAAC 1410
QY 379 LeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPhe 398
Db 1411 CTGCTACCAAGCCCTGGAGCTCCAGCTGAAGCTGAACCCCTGAGCTCCACAGCTGTTT 1470
QY 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnLysValGlnLeu 418
Db 1471 GCCAAGCTGCTCCAGAAATGACAGACCTCCACAGATTTGTACGGAAACAGCTGCAGCTA 1530
QY 419 LeuGlnValIleLysLysThrGlnLysMetSerLeuHisProLeuLeuGlnGluIle 438
Db 1531 CTGCAAGTGATCAGAAAGACGAGACAGACATGAGTCTTCACCCGCTCCTGCAGAGATC 1590
QY 439 TyrLysAspLeuTyr 443
Db 1591 TACAAGACTTGTAC 1605
RESULT 3
AAF23644
ID AAF23644 standard; cDNA; 1608 BP.
XX AAF23644;
AC 27-MAR-2001 (first entry)
XX
DT 27-MAR-2001 (first entry)
XX
DE Human PPARgamma coding sequence.
XX
KW Cytostatic; antiinflammatory; antisense oligonucleotide; PPARgamma;
KW peroxisome proliferator-activated receptor gamma; transcription factor;
KW nuclear hormone receptor; human; infection; inflammation; tumour; ss.
XX
OS Homo sapiens.
XX
PN US6159734-A.
PM 12-DEC-2000.
PD 12-DEC-2000.
PE 18-JAN-2000; 2000US-0484345.
PR 18-JAN-2000; 2000US-0484345.
PS (ISIS-) ISIS PHARM INC.
PI McKay R, Baker BF, Borchers AH;
XX WPI: 2001-070112/08.
DR P-PSDB: AAB59839.
XX Novel antisense compounds capable of modulating expression of
PT peroxisome proliferator-activated receptor gamma useful for diagnosis,
PT prophylaxis and treatment of diseases associated with expression of the
PT receptor -
XX
XX Example 13; Column 45-48; 40pp; English.
XX Peroxisome proliferator-activated receptors (PPARs) are members of the
CC nuclear hormone receptor subfamily of transcription factors. The present
CC invention relates to antisense oligonucleotides (see AAF23652-F23731),
CC targeted to a nucleic acid molecule encoding human PPARgamma, which
CC specifically hybridises with and inhibits the expression of human
CC PPARgamma. The present sequence is the coding sequence for human
CC PPARgamma. The PPARgamma antisense oligonucleotides of the present
CC invention can be used in the diagnosis and treatment of diseases
CC associated with the expression of PPARgamma, e.g. to prevent or delay

CC infection, inflammation or tumour formation.
XX
SQ Sequence 1608 BP; 462 A; 380 C; 361 G; 405 T; 0 other:
Alignment Scores:
Pred. No.: 1,87e-197 Length: 1608
Score: 2231.00 Matches: 443
Percent Similarity: 87.72% Conservative: 0
Best Local Similarity: 87.72% Mismatches: 0
Query Match: 63.60% Indels: 62
Db: 22 Gaps: 1
US-09-931-007a-1 (1-688) x AAF23644 (1-1608)
QY 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20
Db 91 ATGGGTGAACCTCTGGAGATTTCTCTATTGACCCAGAAAGGATTCCTTCACTGATACA 150
QY 21 LeuSerAlaAsnIleSerGlnLysMetThrMetValAspThrGluMetProPheThrPro 40
Db 151 CTGCTGCAAAACATATCAACAAGAAATGACCAATGTTGACACAGAGATGCATTTGGCCC 210
QY 41 ThrAsnPhenGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
Db 211 ACCAACTTTGGGATCAGCTCCGCTGATCTCTCCGAATGGAAGACCACTCCACCTCTTT 270
QY 61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisThrGluAsp 80
Db 271 GATATCAACCCCTTCACTACAGTCTTCACTTCTCCACATTTTCACTCCACATTCAGAAC 330
QY 81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGln 100
Db 331 ATTCATTCACAAAGAACAGATCCAGTTCGACATTCAGATGATGACCTGAAACTTCAA 390
QY 101 GluTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThr 120
Db 391 GAGTACCAAGAGCAATCAAAAGTGAAGCCCTGCATCTCCACCTTAATATTCGAGAAACT 450
QY 121 GlnLeuTyrAsn----- 124
Db 451 CAGCTCTACAAATACCTCATGAAGACCTTCCAACTCCCTCATGGCAATTGAATGTGCT 510
QY 124 ----- 124
Db 511 GTCTGTGAGATAAAGCTTCTGATTTCACTATGAGAGTTCAATGTTGAGAGATGCAAG 570
QY 124 ----- 124
Db 571 GGTTCCTCCGGAACAATCAGATTGAAGCTTATCTATGACAGATGTGATCTTAAGTGT 630
QY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
Db 631 CGGATCCACAAAAGAAAGTAATATGTCTGCTGTGCTTTCAGAAATGCTTGTGA 690
QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158
Db 691 GTGGGATGTCATTAATGCCATCAGTTTGGGCGGATGCCACAGCGGAGAGAGAAAG 750
QY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
Db 751 GTGTTGGCGGAGATCTCCAGTATATGACCACTGGAATCCAGATCCGCTGACCTCCGG 810
QY 179 AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198
Db 811 GCCCTGGCAAAACATTTGTATGATCATACATAAAGTCTTCCCTGACCAAAACAAAG 870
QY 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218
Db 871 GCGAGGCGGATTCATGACGAGAAAGCAACAGACAATACCATTCGTTATCTATGACATG 930
QY 219 AsnSerLeuMetMetGlyLysAspLysIleLysPheLysHisIleThrProLeuGlnGlu 238
Db 931 AATTCCTTAATGATGGAGAAAGATAAATCAAGTTCAAAACATCACCCTCCCTGCAGAG 990

```

QY 239 GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258
    |||
Db 991 CAGACCAAGAGAGTGGCCATCCGCACTTTTCAGGGCTGCGCACTTCCTCCGTGGAGGCT 1050
QY 259 ValGlnGluIleThrGluValAlaLysSerIleProGluPheValAsnLeuAspLeuAsn 278
    |||
Db 1051 GTCCAGAGATCAGACAGATATGCCAAAGCATTTCTGGTTTGTAAATCTTCACTTGAAC 1110
QY 279 AspGlnValThrLeuLeuLysTrpGlyValHisGluIleIleTrpThiMetLeuAlaSer 298
    |||
Db 1111 GACCAAGTAACTCTCTCAAAATATGAGTCCAGATCATTTTACACAAATGCTGGCTCC 1170
QY 299 LeuMetAsnLysAspGluValLeuIleSerGluGlnGlyGlnGlyPheMetThrArgGluPhe 318
    |||
Db 1171 TTGATGATTAAGATGGGGTCTCTCATTCGAGGGCCCAAGCTTATGACAAAGGAGTTT 1230
QY 319 LeuLysSerLeuArgLysProPheGluAspPheMetGluProLysPheGluPheAlaVal 338
    |||
Db 1231 CTAAAGAGCTTCGAAAGCCCTTTTGTGACTTTATGAGACCCCAAGTTTGAGTTTCTGTG 1290
QY 339 LysPheAsnAlaLeuGluLeuAspSerSerAspLeuAlaIlePheIleAlaValIleIle 358
    |||
Db 1291 AAGTTCATGCACTGGAAATGATGACAGCACTTGGCAATATTTATTCCTCATTTATT 1350
QY 359 LeuSerGluAspArgProGluLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
    |||
Db 1351 CTCATGTGAGAGACGCCGCGAGTTTCTGTAATGTGAAGCCCATTTGAAGACATTTAAACAC 1410
QY 379 LeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerGlnLeuPhe 398
    |||
Db 1411 CTGCTACAAAGCCCTGGAGCTCAGCTGAAGTGAACCAACCTGAGTCTCCACAGCTGTTT 1470
QY 399 AlAluLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeu 418
    |||
Db 1471 GCCAAGCTGCTCCAGAAAATGACAGACCTCAGACAGATTTGTACAGGAACACCTGACGCTA 1530
QY 419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIle 438
    |||
Db 1531 CTGCAGGTGATCAAAAGACGAGACACATGACTTTTACCCCTCTGACAGAGATC 1590
QY 439 TyrLysAspLeuTyr 443
    |||
Db 1591 TACAAGGACTGTGTAC 1605

RESULT 4
AAT35334
ID AAT35334 standard; DNA; 1647 BP.
XX
XX AAT35334;
XX AC
XX 14-APR-1997 (first entry)
XX DT
XX 14-APR-1997 (first entry)
XX DE
XX Peroxisome proliferator activated receptor gamma2 coding sequence.
XX KW
XX Peroxisome proliferator activated receptor gamma; hPPARgamma; human;
KW peroxisome; cholesterol metabolism; lipid metabolism; respiration;
KW fatty acid metabolism; beta-oxidation cycle enzyme; hypolipidemic drug;
KW herbicide; hPPARgamma2; adipose tissue; obesity; diabetes; anorexia;
KW hyperlipidaemia; hypercholesterolaemia; hyperlipoproteinaemia; antibody;
KW metabolic disorder; therapy; ss.
XX
XX Homo sapiens.
XX OS
XX
XX Key Location/Qualifiers
XX PH 130..1647
XX FT CDS
XX FT /tag= a
XX FT /product= hPPARgamma2 subtype 1
XX FT 214..1647
XX FT /tag= b
XX FT /product= hPPARgamma2 subtype 2
XX FT 220..1647
XX FT CDS
XX FT /tag= c

```

```

FT /product= hPPARgamma2 subtype 3
XX
XX MO9623884-A2.
XX
XX 08-AUG-1996.
XX
XX 29-JAN-1996; 96WO-US01469.
XX
XX 23-OCT-1995; 95US-0005809.
XX 30-JAN-1995; 95US-0380051.
XX 07-JUN-1995; 95US-0484487.
XX
XX (LIGAND) LIGAND PHARM INC.
XX
XX Mukherjee R;
XX
XX WPI: 1996-384114/38.
XX P-PSDB: AAR99326-R99328.
XX
XX Human peroxisome proliferator activated receptor(s), hPPAR-gamma and
XX hPPAR-gamma2 - useful to identify polypeptide (ant)agonists to treat
XX e.g. obesity, anorexia and diabetes
XX
XX Claim 43: Page 87-88; 108pp; English.
XX
XX This sequence represents the coding sequence for the human peroxisome
XX proliferator activated receptor gamma2 (hPPARgamma2). This sequence
XX encodes three different hPPARgamma2 subtypes, which differ in their
XX N-terminal sequences. Peroxisomes contain enzymes for cholesterol and
XX lipid metabolism, and respiration. Peroxisome proliferators increase the
XX capacity of the peroxisomes to metabolise fatty acids, via increased
XX expression of the beta-oxidation cycle enzymes. Peroxisome proliferators
XX include unsaturated fatty acids, hypolipidemic drugs, and herbicides.
XX hPPARgamma has two main subtypes (hPPARgamma, and hPPARgamma2), which
XX differ in their N-terminal sequences, and are expressed at high levels in
XX adipose tissue. hPPARgamma proteins repress hPPARalpha activity. The
XX encoded proteins can be used in a cell system with a reporter gene (and
XX optionally a hPPARalpha protein) to identify agonists and antagonists of
XX the polypeptides, which can be used as therapeutic agents. The agonists
XX and antagonists can be used to treat obesity, diabetes, anorexia,
XX hyperlipidaemia, hypercholesterolaemia, hyperlipoproteinaemia, and other
XX metabolic disorders. This sequence, the encoded proteins, and antibodies
XX against them, can be used to establish a tissue specific expression
XX pattern for hPPARgamma.
XX
XX SO Sequence 1647 BP; 472 A; 389 C; 371 G; 415 T; 0 other:

Alignment Scores:
Pred. No.: 1.93e-197 Length: 1647
Score: 2231.00 Matches: 443
Percent Similarity: 87.72% Conservative: 0
Best Local Similarity: 87.72% Mismatches: 0
Query Match: 63.60% Indels: 62
DB: 17 Gaps: 1

US-09-931-007A-1 (1-688) x AAT35334 (1-1647)
QY 1 MetGlyGluThrLeuGlnLysSerProIleAspProGluSerAspSerPheThrAspThr 20
    |||
Db 130 ATGGGTGAACCTCGGAGATTTCTTGTGACCCAGAAACGATTCCTTCACTGATACA 189
QY 21 LeuSerAlaAsnLysSerGlnGluMetThrMetValAspThrGluMetProPheThrPro 40
    |||
Db 190 CTGCTGCAAAACATATTCACAGAAATGACCATGTGTTGACACAGAGATGCCATTTGGCCC 249
QY 41 ThrAsnPheGlyLysSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
    |||
Db 250 ACCAAGTTTGGGATCAGCTCCGTGATCTCCGTAATGGAAGACATCCCATCTTT 309
QY 61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80
    |||
Db 310 GATATCAAGCCCTTCACATGATGTGACTTTCACAGCATTTCTACTCCACATTAAGAAAGAC 369

```

OY	81	ILPRophethrAargThrAspProValValAlaAspTylsTyrAspleuNlsleuGln	100
Db	370	ATTCATTCCACAAACAGATCCAGTGGTTGCAGAGATTACAAAGTATCACTGGAACCTTCAA	429
OY	101	GIuTyrGlnSerAlaIleLysValGlnProAlaSerProforTyrIserGluTyrThr	120
Db	430	GAGTACCAAGTGCATATCAAGTGGACCTGCATCTCCATTATTATTCGTGAGAGACT	489
OY	121	GlnLeuTyrAsn-----	124
Db	490	CAGCTCTACAAATAAGCCATGAGAGACCTTCCAACTCCCTCATGGCCAAATGTGATGTCT	549
OY	124	-----	124
Db	550	GTCGTGGAGATAAAGCTTTCGATTTCACTATGAGATTTCATGCTGTGAAGATGCAAG	609
OY	124	-----	124
Db	610	GGTTTCTTCCGGACAAACATCAGATTGACCTTATGACAGATGATCTTAACTGT	669
OY	125	-----ArgAsnLysCysGlnTyrCysArpPheGlnLysCysLeuAla	138
Db	670	CGGATCCACAAAAAGAGAAATAATATGTCAGTACGTGCGTTTCAGAAATGCCCTGGA	729
OY	139	ValGlyMetSerHisAsnAlaIleArgPheGlyArgmetProGlnAlaGluTsgLys	158
Db	730	GTTGGGAGGTCTCAATATGCCATCAGTGGTGGCGGATGTCACAGGCCAGGAAGGAGAG	789
OY	159	LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAspProGlnSerAlaAspleuArg	178
Db	790	CTGTGGCGGGAGATCTCAGTGAATTCAGCCACTGATATCCAGATGCCCTGACCTCGG	849
OY	179	AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys	198
Db	850	GCCCTGGCAAAACATTTGTATGATCATCATATAAGTCCITCCCGCTGACCAAGCAAG	909
OY	199	AlaArgAlaIleLeuThrGlyLysThrAspLysSerProPheValIleTyrAspMet	218
Db	910	GCGAGCGGACCTTGACACGGAACACAAACAAATCACCATTCCTTATCATGACATG	969
OY	219	AsnSerLeuMetMetGlyGluAspLysIleLysPheLysHisIlePheProLeuGlnGlu	238
Db	970	AATTCCTTAATGATGGAGAGAGATAAATCAAGTCAACACACTCACCCCTGCAGGAG	1029
OY	239	GlnSerLysGluValAlaIleArgGlyIlePheGlnGlyCysGlnPheArgSerValGluAla	258
Db	1030	CAGAGCAAAAGGTGGCCATCCGATCTTCAGAGCGTCCGACATTTGCTCCGTGCAGCT	1089
OY	259	ValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspleuAsn	278
Db	1090	GTCGAGGAGATCACAGAGTATGCCAAAGCATTTCCGTGTTGTAAATCTTGACTTGAAC	1149
OY	279	AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer	298
Db	1150	GACCAAGTAACTCTCTCCAAATATGGAGTCCAGAAATCAATTACCAATAGCTGCCTCC	1209
OY	299	LeuMetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGluPhe	318
Db	1210	TGTATGATATAAAGATGGGGTTCATATCCGAGGCGCAAGGCTCATGCAAGGAGATTT	1269
OY	319	LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal	338
Db	1270	CTAAGAGGCTGCGCAAAAGCCCTTTGGTATGACTTATGAGCCCAAGTTTGAAGTTTCTGTG	1329
OY	339	LysPheAsnAlaLeuGluLeuAspAspSerAspLeuAlaIlePheIleAlaValIleLe	358
Db	1330	AAGTTCAATGACACGGAAATTAAGATGACGAGCATTTGCAATATTATTCCTGCTATTAT	1389
OY	359	LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspSn	378
Db	1390	CTCAGTGGAGACGCCCGCAGGTTTGCCTGAATGTGAACCCATTGACACATTTGAACACAC	1449
OY	379	LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluLysSerGlnLeuPhe	398

ID	Sequence	Location/Qualifiers
Db 1450	CTGCTACCAAGCCCTGGAGCTCCACCTGAAACCTGAACACCCTGAGTCTCCACAGCTGTTT	
Qy 399	AlaIysIleuGlnIysMetThrAspLeuArgGlnIleValThrGlnIysValGlnIleu	
Db 1510	GCCAAAGCTGCTCCGAAATAGTACAGACCTTAGACACATTTGTACAGGACACCTGCAGCTA	
Qy 419	LeuGlnValIleIysIysThrGlnThrAspMetSerLeuGlnIleProIleuGlnIle 438	
Db 1570	CTGCAGGTGATCAGAGAGAGGAGACAGACATGATGCTTACCCCGCTCTCAGAGAGATC	
Qy 439	TyrlsAspLeuTyr 443	
Db 1630	TACAGAGACTGTGTAC 1644	
RESULT 5		
AAH76294		
ID	AAH76294	standard; cDNA: 2596 BP.
XX	AAH76294:	
AC		
XX		
DT	29-OCT-2001	(first entry)
XX		
DE	Human PAX8e9(-exon 8)-PARgammae1 cDNA sequence.	
XX		
XX	PAX8-PPARGammae1; oncogene; cytostatic; PAX8; PPARGammae1; cancer;	
KW	follicular carcinoma; PAX8e9(-exon 8)-PPARGammae1; human; ss.	
XX		
OS	Homo sapiens.	
XX		
FT	Key	Location/Qualifiers
FT	CDS	161..2596
FT		/*tag= a
PN	WO200152789-A2.	
XX		
XX	26-JUL-2001.	
PD		
XX		
PF	18-JAN-2001; 2001WO-US01664.	
XX		
PR	20-JAN-2000; 2000US-0177109.	
XX	14-AUG-2000; 2000US-0225079.	
PA	(BGHM) BRIGHAM & WOMENS HOSPITAL INC.	
PI	Kroll TG, Fletcher JA;	
XX	WPI: 2001-514487/56.	
DR	P-PSDB: AAB85801.	
XX		
PT	New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and	
XX	treating certain tumors or cancers, e.g. follicular carcinoma -	
PS	Claim 1; Page 127-131, 145pp: English.	
XX		
CC	The invention relates to an oncogene designated PAX8-PPARGammae1 that	
CC	contains a PAX8 coding region fused to PPARGammae1 coding region. The	
CC	PAX8-PPARGammae1 polypeptides can be expressed by standard recombinant	
CC	methodology. A PPARGammae1 ligand or agent is useful for treating a	
CC	subject having a disorder characterized by the presence of a PAX8-	
CC	PPARGammae1, where the disorder is cancer, e.g. follicular carcinoma.	
CC	The PAX8-PPARGammae1 molecules are also useful for providing nucleotide	
CC	and amino acid sequences useful for detecting the above disease. The	
CC	present sequence represents a human PAX8e9(-exon 8)-PPARGammae1	
XX	polypeptide encoding cDNA.	
Sequence	2596 BP; 660 A; 745 C; 672 G; 519 T; 0 other;	
Alignment Scores:		
Pred. No.:	1 23e-185	Length: 2596
Score:	2106.50	Matches: 425
Percent Similarity:	84.45%	Conservative: 4
Percent Local Similarity:	83.66%	Mismatches: 14

Query Match:	60.05%	Indels:	65
DB:	22	Gaps:	3
US-09-931-007A-1 (1-688) x AAH76294 (1-2596)			
QY	1 MetGIgIuHrIleu---GIyAspSerProIleAspProGluSer-----AspSerPhe	17	
Db	1070 GTGGGGCCACAGCTGCCCCGATATCCACCCACATCCACAGGGGACAGGACGCTAT	1129	
QY	18 ThrAspThrIleuSerAlaAsnIleSerGlnIuMetIhrMetValAspThrGluMetPro	37	
Db	1130 GCCTCTCTGTCATCGCAGGACGATGTGGCAAAATGACCATGTTGACACAGAGATGCCA	1189	
QY	38 PheTrpProThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSer	57	
Db	1190 TTCGGGCCACCAACTTGGGATCAGCTCCGATGATCTCCGTAATGGAAGACCACACC	1249	
QY	58 HisSerPheAspIleuSerProPheThrTrpValAspPheSerSerIleSerThrProHis	77	
Db	1230 CACTCCCTTGTATATCAAGCCCTTCACTACTGTTGACTTCTCCAGCATTTCTACTCCAT	1309	
QY	78 TyrGluAspIleProPheThrArgThrAspProValAlaIleAspTyrLysTyrAspLeu	97	
Db	1310 TACGAAGACATTCATTCACAAAGACAGATCCAGTGTTGCAGATTTACAGATATGACCTG	1369	
QY	98 LysLeuGlnIuTyrGlnSerAlaIleLysValGluProAlaSerProTofTyrTyrSer	117	
Db	1370 AACTTCAAGAGTACCAAAATGCAATCAAAAGGAGGCCGTCATCCACCTTATATATCT	1429	
QY	118 GluIuThrGlnIeuTyrAsn	124	
Db	1430 GAGAAGACTCAGCTCATACATATAGCCTCATGAAGACCTTCCAACTCCCTATGGCATTT	1489	
QY	124 -----	124	
Db	1490 GAATGTCGTCTGTGAGATAAAGCTTCTGGATTTCACATGAGATTCAAGCTTGTGAA	1549	
QY	124 -----	124	
Db	1550 GGATGCAAGGTTTCTTCGGAGACATACAGTTGAACTTATCTATGACAGATGTGAT	1609	
QY	125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLys	135	
Db	1610 CTTAACGTGCGATCCACAAAANAAGTAGAANAATATGTCATCTACTGCGTTTCAGAAA	1669	
QY	136 CysLeuAlaValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGlu	155	
Db	1670 TGCCTTCACTGGGAGATGTCATATATGTCATACAGATTTGGCGGATCCACAGCGGAG	1729	
QY	156 LysGluLysLeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAla	175	
Db	1730 AAGGAGAAGCTGTTGGCGGAGATCTCCATGTGATATGACAGACCTAAATCCAGAGTCGCT	1789	
QY	176 AspLeuArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThr	195	
Db	1790 GACCTCGGGCCCTGGCAAAACATTTGTATGACTCATATAAATCTCTCCGCTGACC	1849	
QY	196 LysAlaLysAlaArgAlaIleLeuThrGlyLysTrpThrAspLysSerProPheValIle	215	
Db	1850 AAAACCAAGGCGAGGGCGATCTTTCAGACAGAAAGACACAGACAAATCCACATTCGTTATC	1909	
QY	216 TyrAspMetAsnSerLeuMetGlyGluAspLysIleLysPheLysHisIleThrPro	235	
Db	1910 TATACATGTAATCTTATATGATGGAGAGATTAATTAATCAAGTTCAACACATACACCCC	1969	
QY	236 LeuGlnGlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSer	255	
Db	1970 CTCGAGAGACAGACAAAGAGGTGGCCATCCCATCTTTCAGAGGCTGCCATTTCCGCTCC	2029	
QY	256 ValGluAlaValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeu	275	
Db	2030 GTGGAGGCTCTGAGAGATCACAGATATGTCAAAAGCATTTCCGTTTGTAAATCTT	2089	

QY	276	AspleuasnaspGlnValThrIleuLeuLysTyrGlyValHisGluIleIleTyrThrMet	295
Db	2090	GACTTGAACGCGCAAGTAACTCTCCAAATATGAGAGTCCAGAGATCATTTATACCAATG	2149
QY	296	LeuAlasertLeuMeTAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThr	315
Db	2150	CTGGCCTTCCTTGATGATTAAGATGGGGTTCTCATATCCGAGGGCCAGGCTTCATGACA	2209
QY	316	ArgGluPheLeuLysSerLeuArgLysProPheGlyAspPheMetClnProLysPheGlu	335
Db	2210	AGGGAGTTTCTAAAGAGCCTCGCGAAAGCCTTTGGTGACTTTATGAGAGCCCAAGTTTGAG	2269
QY	336	PheAlaValLysPheAsnAlaLeuGluLeuAspAspSerAspLeuAlaIlePheIleAla	355
Db	2270	TTTTGCTGTGAAGTTCAATGCACCTGGAATTAGATGACAGCGACTTGGAATATTATTGCT	2329
QY	356	ValIleIleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGlnAspIle	375
Db	2330	GTCATTATTCTCACTGGAGAGCCGCCAGGTGCTGAAATGTAAGCCCAATTGAACACTT	2389
QY	376	GlnAspAsnLeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerSer	395
Db	2390	CAAGCAACACCGCTACACAGCCCTCGAGCTCCAGCTGAACCTGACCACCTGAGTCTCA	2449
QY	396	GlnLeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHis	415
Db	2450	CAGCTGTTTGCCAGAGCTGCTCCAGAAATGACAGACCTCAGACAGATTGTCCAGGACAC	2509
QY	416	ValGlnLeuLeuGlnValIleIleLysLysThrGluThrAspMetSerLeuHisProLeuLeu	435
Db	2510	GTCGACGCTACTGCGAGTGATCAAGAAGACGAGACAGACATGATGATCTTACCCGCTCTG	2569
QY	436	GlnClnIleTyrLysAspLeuTyr	443
Db	2570	CAGGAGATCTACAGAGCTGTGAC	2593
RESULT 6			
ID	AAH76283	standard; cDNA; 2625 BP.	
XX	AAH76283;		
AC			
DT	29-OCT-2001	(first entry)	
XX			
DE	Human PAX8e9-PPARGammae1 cDNA sequence.		
XX			
KW	PAX8-PPARGamma1; oncogene; cytostatic; PAX8; PPARGamma1; cancer;		
KW	follicular carcinoma; PAX8e9-PPARGammae1; human; ss.		
OS	Homo sapiens.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	1..2625	
FT	/*tag= a		
XX	WO200152789-A2.		
PN			
XX	26-JUL-2001.		
PD			
XX			
PF	18-JAN-2001; 2001MO-US01664.		
XX			
PR	20-JAN-2000; 2000US-0177109.		
PR	14-AUG-2000; 2000US-0225079.		
XX			
PA	(BGHM) BRIGHAM & WOMENS HOSPITAL INC.		
XX			
P1	Kioli TG, Fletcher JA;		
XX			
DR	WPI: 2001-51487/56.		
DR	P-PSDB: AAB85795.		
XX			
PT	New PAX8-PPARGc1 oncogene and oncoprotein, useful for detecting and		
PT	treating certain tumors or cancers, e.g. follicular carcinoma -		

XX Claim 1: Page 112-116; 145pp; English.

XX
XX
CC The invention relates to an oncogene designated PAX8-PPARgamma1 that
CC contains a PAX8 coding region fused to PPARgamma1 coding region. The
CC PAX8-PPARgamma1 polypeptides can be expressed by standard recombinant
CC methodology. A PPARgamma1 ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARgamma1, where the disorder is cancer, e.g., follicular carcinoma.
CC The PAX8-PPARgamma1 molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The
CC present sequence represents a human PAX89-PPARgamma1 polypeptide
CC encoding cDNA.

XX
SQ Sequence 2625 BP; 662 A; 761 C; 646 G; 556 T; 0 other;

Alignment Scores:

Pred. No.: 1,25e-185 Length: 2625
Score: 2106.50 Matches: 425
Percent Similarity: 84.45% Conservative: 4
Best Local Similarity: 83.66% Mismatches: 14
Query Match: 60.05% Indels: 65
DB: 22 Gaps: 3

US-09-931-007A-1 (1-688) x AAH76283 (1-2625)

QY 1 MetGlyGlnThrLeu---GlyAspSerProIleAspProGluSer-----AspSerPhe 17
Db 1099 GTGGGGCCCAACGCTGCCCGGATACCCACCCACATCCCCACACGACGACAGGCGACTAT 1158
QY 18 ThrAspThrLeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetPro 37
Db 1159 GCCTCCCTCCATCCGACGATGCTGGCAGAAATGACCATGCTGACACAGAGATGCCA 1218
QY 38 PheTrpProThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSer 57
Db 1219 TTTGGGCCCAACACTTGGATGAGTCCGCGATCTCTCCGAAAGGAAGACACTCC 1278
QY 58 HisSerPheAspIleLysProPheThrThrValAspPheSerSerIleSerThrProHis 77
Db 1279 CACTCCCTTGATATCAAGCCCTTCACACTGCTGACTTCCACCATTTCTACTCCACAT 1338
QY 78 TyrGluAspIleProPheThrArGThrAspProValAlaAlaAspTyrIleTyrAspLeu 97
Db 1339 TAGAGAGACATTCATTCACAGAACAGATCCAGTGGTGCAGATTCACAGATAGACCTG 1398
QY 98 LysLeuGlnGluTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSer 117
Db 1399 AAACCTCAAGAGTACCAAAAGTGAATCAAAAGTGAAGCTCATCTCCACTTATTATTC 1458
QY 118 GlnLysThrGlnLeuTyrAsn----- 124
Db 1459 GAGAAGCTCAGCTTACATTAAGCCCTCATGAGAGCCTTCCAACCTCCATAGGCAATT 1518
QY 124 ----- 124
Db 1519 GAATGTCGTCTGTGAGAGATAAAGCTTCTGATTCATATGAGATTCATGCTTGCAA 1578
QY 124 ----- 124
Db 1579 GGATGCAGAGGGTTCTTCGCGAGAACATCAGATTGAAGCTTATATGACAGATGTGAT 1638
QY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLys 135
Db 1639 CTTAAGTGTGATCCACAAAAAAGTAGAATAATATGATAGTACTGTGCTTTCAGAAA 1698
QY 136 CysLeuAlaValAlaLysMetSerHisAsnAlaIleArgPheGlnTyrMetProGlnAlaGlu 155
Db 1699 TGCCTTGCAATGGGAGTGTCTCATTAATGACATCAGGTTTGGCGGAGTGCACAGCGCAG 1758
QY 156 LysGlnLysLeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAla 175
Db 1759 AAGGAGAAGCTGTTGGCGGAGATCTCCAGATATGACACAGCTGAATCCAGAGTCCGCT 1818

QY 176 AspLeuArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThr 195
Db 1819 GACCTCCGGCCCTGGCAAAACATTTTGTATGACTCATACATTAAGTCTCTCCGCTGACC 1878
QY 196 LysAlaLysAlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIle 215
Db 1879 AAACAAAGCGGAGCGGATCTTGACAGAAAGACACACAAATCACCATTCCGTTATC 1938
QY 216 TyrAspMetLysSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrPro 235
Db 1939 TATGACATGAAATCTTAAAGATGGGAGAAATAAATCAAGTTCAAAACATCACCCCC 1998
QY 236 LeuGlnGlnLeuSerLysGluValAlaIleArgIlePheGlnGlnLysGlnPheArgSer 255
Db 1999 CTCGAGAGCAGACAGCAAAAGAGTGGCCATCCGATCTTTAGAGGCTCCAGTTCCGCTCC 2058
QY 256 ValGluAlaValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeu 275
Db 2059 GTGGAGGCTGTGCAGGATCAGATCAGAGTATGCCAAAGCATTCCTGGTTTGTAAATCTT 2118
QY 276 AspLeuAsnAspGlnValIleThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMet 295
Db 2119 GACTTGAACGACCAAGTAACCTCTCAATATGGAATGGAATCCACAGATCATTTACACATG 2178
QY 296 LeuAlaSerLeuMetLysAsnLysAspGlyValLeuIleSerGlnGlnGlnLysPheMetThr 315
Db 2179 CTGGCCCTCTTGATGAATTAAGATGGGTTCTCATATCCAGAGGCCCAAGCCTTCATGACA 2238
QY 316 ArgGluPheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGlu 335
Db 2239 AGGAGATTTCTAAAGAGCCCTGCAGAAAGCCCTTGTGACTTATATGAGCCCAAGTTTGAG 2298
QY 336 PheAlaValLysPheAsnAlaLeuGluLeuAspAspSerLysPheLeuAlaIlePheIleAla 355
Db 2299 TTTGCTGTGAAGTTCATGACCTGGAATTAGATGACAGCCGATTTGGCAATATTTATTTGCT 2358
QY 356 ValIleLeuLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIle 375
Db 2359 GTCATTTTCTCAGTGTGAGAGCCGCCAGCTTCTGATGATGAAAGCCCATTTGAAGACATT 2418
QY 376 GlnAspAsnLeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisPheGluSer 395
Db 2419 CAAGACAACTGCTTACAGGCCCTGGAGCTCCAGCTGAAGCTGAACCAACCTGAGTCTCA 2478
QY 396 GlnLeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValIleThrGlnHis 415
Db 2479 CAGCTGTTTGCAGAGCTGCTCCAGAAATGACAGACCTCAGACAGATTCACAGGAACAC 2538
QY 416 ValGlnLeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisPheLeuLeu 435
Db 2539 GTGCAGCTACTGACAGTGATCAAGAAAGACGACGACAGACATGATGATCTTCCACCGCTCTG 2598
QY 436 GlnGluIleTyrLysAspLeuTyr 443
Db 2599 CAGGAGATCTACAAAGGACTGTAC 2622
RESULT 7
ID AAH76281 standard; cDNA: 2334 BP.
AC AAH76281;
XX
DT 29-OCT-2001 (first entry)
XX
DE Human PAX89-PPARgamma1 cDNA sequence.
XX
KW PAX8-PPARgamma1; oncogene; cytosolic; PAX8; PPARgamma1; cancer;
KW follicular carcinoma; PAX89-PPARgamma1; human; ss.
XX
OS Homo sapiens.
OS
FH Key Location/Qualifiers

FT CDS 1..2334
FT /*tag= a
XX MO200152789-A2.
XX 26-JUL-2001.
XX 18-JAN-2001; 2001WO-US01664.
XX 20-JAN-2000; 2000US-0177109.
XX 14-AUG-2000; 2000US-0225079.
XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.
XX PA
XX Kroll TG, Fletcher JA;
XX WPI: 2001-514487/56.
XX DR P-PSDB: AAB85793.
XX
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
XX treating certain tumors or cancers, e.g. follicular carcinoma -
XX
XX Claim 1: Page 100-104; 145pp; English.
XX
XX The invention relates to an oncogene designated PAX8-PPARgamma1 that
XX contains a PAX8 coding region fused to PPARgamma1 coding region. The
XX PAX8-PPARgamma1 polypeptides can be expressed by standard recombinant
XX methodology. A PPARgamma1 ligand or agent is useful for treating a
XX subject having a disorder characterized by the presence of a PAX8-
XX PPARgamma1, where the disorder is cancer, e.g. follicular carcinoma.
XX The PAX8-PPARgamma1 molecules are also useful for providing nucleotide
XX and amino acid sequences useful for detecting the above disease. The
XX present sequence represents a human PAX8e7-PPARgamma1 polypeptide
XX encoding cDNA.
XX
XX Sequence 2334 BP: 614 A; 650 C; 573 G; 497 T; 0 other:
SQ
Alignment Scores:
Pred. No.: 5,34e-185 Length: 2334
Score: 2099.00 Matches: 420
Percent Similarity: 84.33% Conservative: 5
Best Local Similarity: 83.33% Mismatches: 13
Query Match: 59.83% Indels: 66
DB: 22 Gaps: 2
US-09-931-007a-1 (1-688) x AAB76281 (1-2334)
QY 6 G1yAsPSePrIoIleAsPProGluSeRAsPSePhetHrAsPThLeuSeRa1AAsn--- 24
DB 820 GGGAGAGCCACCCCTGACCCCTCCACACGCCACTGGGGCGCACCTCTCGACTCACCAG 879
QY 25 -----IleSerGlnGluMetHrMetValAsPThrGluMeProPhetProthr 41
DB 880 ACCTAACCCCGTGGGCGAAATGACCATGTGTGACACAGAGATCCATTCTGGCCACC 939
QY 42 AsnPhGlyIleSerSerValAsPLeuSerValMetGluAsPHisSerHisSerPheAsP 61
DB 940 AACCTTGGATGACGTCGCGTGAATCTCTCCGTAATGGAAGACCACTCCACCTCTTGAT 999
QY 62 IleuSProPheThrThrValAsPHeSerSerIleSerThrProHisTyrGluAsP 81
DB 1000 ATCAAGCCCTTCACACTGTTGACTTCTCCAGCATTTCTACTCCACATTACGAAGACATT 1059
QY 82 ProPheThrArGTThrAsPProGluValAlaAsPThrLysTyrAsPLeuLysGlnGlu 101
DB 1060 CCATTTCACAGAGACGATCCATGCTGTGAGATTACAAAGTATGACCTGAAACTTCAAGAG 1119
QY 102 TyrGlnSerAlaIleLysValGluProAlaSerProTyrTyrSerGluLysThrGln 121
DB 1120 TACCAAGTGCATCAAGTGAAGGCTGCATCTCCACCTTATTATTCTGAGAGAGCTCAG 1179
QY 122 LeuTyrAsn----- 124
|||||

DB 1180 CTCTACAAATAAGCTCTAGAAAGGCTTCCAACTCCCTCATGGCAATTGAATGTCGTCTC 1239
QY 124 ----- 124
DB 1240 TGTGAGATAAAGCTTCTGGATTTCTACTATGAGATTCTGCTTGTGAAGATGCCAAGGCT 1299
QY 124 ----- 124
DB 1300 TTCTCCGAGACAAATCAGATTGAAGCTTATCTATGACAGATGATCTTAACGTGCGG 1359
QY 125 -----ArgAsnLysCysGlnTyrCysArGpHeGlnLysCysLeuAlaVal 139
DB 1360 ATCCACAAAAAAGTAGAAATAAATGTCAGTACTGTCGGTTTCAGAAATGGCTTGCACTG 1419
QY 140 GlyMetSerHisAsnAlaIleArGpHeGlyArGpMetProGlnAlaGluLysGluLysLeu 159
DB 1420 GGGATGCTCATTAATGCCATCAGGTTTGGGGGAGTCCACAGGCCGAGCAAGAGAGAGCTG 1479
QY 160 LeuAlaGluIleSerSerAsPLeuAsPLeuAsnProGluSerAlaAsPLeuArGla 179
DB 1480 TTGGCGGAGATCTCCAGTCATATGACAGCTGAAATCCAGACTCCGCTGACCTCGGGCC 1539
QY 180 LeuAlaLysHisLeuTyrAsPSeTyrIleLysSerPheProLeuThrLysAlaLysAla 199
DB 1540 CTGGCAAAACATTTGTATGACTCATATCAATAAGTCTTCCGCTGACCAAGCAAGAGCG 1599
QY 200 ArgAlaIleLeuThrGlyLysThrThrAsPLeuSerProPheValIleTyrAsPMetLsn 219
DB 1600 AGGGGATCTTGACGAGGAAGCAACACAAATCCATTCCTGTTATATATGACCTGAT 1659
QY 220 SerLeuMetGlyLysAsPLeuLysIleLysPheLysHisIleThrProLeuGlnGluGln 239
DB 1660 TCCTTAATGATGGGAGGAATGAATCAAGTTCAACACATCACCCCTCGAGAGAGCAG 1719
QY 240 SerLysGluValAlaIleArGlePheGlnGlyCysGlnPheArGSerValGluAlaVal 259
DB 1720 AGCAAAAGAGTGGCCATCCGATCTTTCAGGGCTCCAGTTCCGCTCGTGGAGGCTGTG 1779
QY 260 GlnGluIleThrGlnTyrAlaLysSerIleProGluPheValAsnAsPLeuAsnSp 279
DB 1780 CAGGAGATCACAGATATGACCAAAAGCATCTCTGTTTGTAAATCTTGACTGACCTGACAGAC 1839
QY 280 GlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeu 299
DB 1840 CAAGTAACCTCCCTCAATATGAGACTCCAGCAGATCATTTACACATCTGGCTCCCTTG 1899
QY 300 MetAsnLysAsPLeuValLeuIleSerGlnGlyGlnIlePheMetThrArgGluPheLeu 319
DB 1900 ATGAATTAAGATGGGCTTCTCATATCCAGAGGCCAAGGCTTCATGACAAAGAGATTCTTA 1959
QY 320 LysSerLeuArGlySProPheGlyAsPHeMetGluProLysPheGluPheAlaValLys 339
DB 1960 AAGACCTCGCCAAACCTTTGGTACTTTATGAGGCCCAAGTTGATGATTTGCTGTGAAG 2019
QY 340 PheAsnAlaLeuGluLeuAsPSeSerAsPLeuAlaIlePheIleAlaValIleLeu 359
DB 2020 TTCAATGACACTGGAATTGATGACAGCAGCTTGGCAATATTATTCGCTCATTAATTCTC 2079
QY 360 SerGlyAsPArGProGlyLeuLeuAsnValLysProIleGluAsPLeuIleGlnAsPAsnLeu 379
DB 2080 AGTGAGAGACCCCGAGGTTTCTGTAATGTGAAGCCCATTTGAAGACATTCAGACAACTG 2139
QY 380 LeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPheAla 399
DB 2140 CTACAGCCCTGAGACTCAGCTGAAGCTGAACCACTCTGAGTCTCCACACTGTTTCC 2199
QY 400 LysLeuLeuGlnLysMetThrAsPLeuArGlnIleValIleThrGlnHisValGlnLeuLeu 419
DB 2200 AAGCTGCTCCAGAAATATACAGACTCAGACAGATTTGACAGGAACAGCTGACAGCTACTG 2259
QY 420 GlnValIleLysLysThrGluThrAsPMetSerLeuHisProLeuLeuGlnGluIleTyr 439
DB 2260 CAGGTGATCAAGAAAGACGAGACAGACATGAGTCTTCCACCGCTCTGACAGAGATCTAC 2319
|||||


```
QY 246 ArgIlePheGlnGlyCysGlnPheArgSerValGluAlaValGlnGluIleThrGluTyr 265
|||||
Db 1045 CCACACTTTTCAGAGGCTGCCAGATTTCGCTCCGTGGAGGCTGTGCAGAGATTCACAGAGTAT 1104
QY 266 AlaIlySerIleProGlyPheValAsnLeuAspLeuAsnAspGluValThrLeuLeuLys 285
|||||
Db 1105 GCCAAAGCATTCCTGGTTTGTAAATCTTGACTGAAAGCAGCAAGTAACTCTCCACAA 1164
QY 286 TTYGIVAlHISGluIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGluVal 305
|||||
Db 1165 TATGAGATCCACGAGATCATTTTACACAAATGCTGGCTCTCTTATGATTAAGATGGGTT 1224
QY 306 LeuIleSerGluGlyGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysPro 325
|||||
Db 1225 CTCATATCCGAGGCGCCAGGCTTCATGACAAAGAGTTTCTAAAGACCTGGCAAGACCT 1284
QY 326 PheGlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeu 345
|||||
Db 1285 TTYTGCTGACTTATGAGAGCCCAAGTTTGAGTTTGCTGTGAAATGCAATGCACTGGAATTA 1344
QY 346 AspAspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGly 365
|||||
Db 1345 GATGACAGGACTGGCAATATTATGCTGTGATTAATTCAGTGGAGACCGCCAGGT 1404
QY 366 LeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGluLeu 385
|||||
Db 1405 TTGCTGAATGTGAAGCCCATTTGAAGCATTTCAAGACAACTGTCTACAAACCTGGAGCTC 1464
QY 386 GlnLeuLysLeuAsnHisProGluSerSerGlnLeuPheAlaLysLeuGlnLysMet 405
|||||
Db 1465 CAGCTGAACTGACACCCCTGAGTCTCCACAGCTGTTGCCAAGCTGCTCCAGAAATG 1524
QY 406 ThrAspLeuArgGlnIleValThrGluHisValGlnLeuGlnValIleLysLysThr 425
|||||
Db 1525 ACAGACTCTGACAGATTGTCTCAGGACACGTCAGCTACTCGAGGTGATCAAGAGACG 1584
QY 426 GluThrAspMetSerLeuHisProLeuLeuGlnGluIleTyrLysAspLeuTyr 443
|||||
Db 1585 GAGACAGACATGATGCTTACCCGCTCGTGCAGAGATCTACAGACTTGTACTAGCAG 1644
QY 443 ----- 443
Db 1645 AGAGTCTGAGCCACTGCCACATTTCCCTTCTCCAGTTGCACTATTCTGAGCCGGAAT 1704
QY 444 AlaIlePheIleLeuThrGlyLysThrThrAspLysSer 456
|||||
Db 1705 TCTTTTGTCTTTTACCCTGGAAAGAAATCTCATTAAGC 1743
RESULT 9
AAH76282
ID AAH76282 standard; cDNA; 2523 BP.
XX
AC AAH76282;
XX
DT 29-OCT-2001 (first entry)
XX
DE Human PAX8e8-PPARGammae1 cDNA sequence.
XX
KM PAX8-PPARGammae1; oncogene; cytosolic; PAX8; PPARGammae1; cancer;
KW Follicular carcinoma; PAX8e8-PPARGammae1; human; ss.
OS Homo sapiens.
XX
FH Key 1.2523 Location/Qualifiers
FT CDS /tag= a
XX
XX WO200152789-A2.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US01664.
XX
```

```
PR 20-JAN-2000; 2000US-0177109.
PR 14-AUG-2000; 2000US-0225079.
XX
XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
PA Kroll TG, Fletcher JA;
XX
XX MPI: 2001-514487/56.
DR P-PSDB; AAB85794.
XX
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma
XX
PS Claim 1; Page 106-109; 145pp; English.
XX
CC The invention relates to an oncogene designated PAX8-PPARGamma1 that
CC contains a PAX8 coding region fused to PPARGamma1 coding region. The
CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant
CC methodology. A PPARGamma1 ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.
CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The
CC present sequence represents a human PAX8e8-PPARGammae1 polypeptide
CC encoding cDNA.
SQ Sequence 2523 BP; 643 A; 723 C; 614 G; 543 T; 0 other:
Alignment Scores:
Pred. No.: 1,72e-184 Length: 2523
Score: 2094.00 Matches: 416
Percent Similarity: 87.03% Conservative: 0
Best Local Similarity: 87.03% Mismatches: 62
Query Match: 59.69% Indels: 1
Gaps: 1
US-09-931-007A-1 (1-688) x AAH76282 (1-2523)
QY 28 GluMetThrMetValAspThrGluMetProPheThrProThrAsnPheGlyIleSerSer 47
|||||
Db 1087 GAATGACCATGATGTGACACAGATGCAATTCCTG6CCACCACTTGGGATCAGCTCC 1146
QY 48 ValAspLeuSerValMetGluAspHisSerHisSerPheAspIleLysProPheThrThr 67
|||||
Db 1147 GTGCAATCTCCGTATGAGAAACCACTCCACTCTTGATATCAAGCCCTCACTACT 1206
QY 68 ValAspPheSerSerIleSerThrProHisTyrGluAspIleProPheThrArgThrAsp 87
|||||
Db 1207 GTTGACTTCTCCAGCATTTCTACTCCACTTACGAAGACATTCATTCACAAGAACAGAT 1266
QY 88 ProValValAlaAspTyrTyrAspLeuLysLeuGlnGluTyrGlnSerAlaIleLys 107
|||||
Db 1267 CCAGTGGTTGCCAGATTATACAGATGACCTGAAACTTCAAGATACCAAGTCAATCAAA 1326
QY 108 ValGluProAlaSerProProTyrTyrSerGluLysThrGlnLeuTyrAsn 124
|||||
Db 1327 GTGAGACCGCATCTCCACTTATTTATTTGTGAGAGACACACTCTACATATAGCCTCAT 1386
QY 124 ----- 124
Db 1387 GAAGAGCCTTCCAACTCCCTCATGCAATTGAATGTCGTGTGAGATTAAGCTTCT 1446
QY 124 ----- 124
Db 1447 GGATTTCATATGAGATTCATGCTGTGAAGATGCAAGGGTTTCTTCGGGAAACAATC 1506
QY 125 -----Arg 125
Db 1507 AGATTGAAGCTTATCTATGACAGATGTGATCTTAATCTGCGGATCCACAAAAAAGGTACA 1566
QY 126 AsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAla 145
|||||
Db 1567 AATAAATGTCACTACTGTCGTTTCAGAAATGCTTGTGAGTGGGATGTCTCATATATGCC 1626
```

```
QY 146 ILeaRgPheGlyArGMeTProGInAlaGluLysGluLysLeuLeuAlaGluIleSerSer 165
XX |||
OS Homo sapiens.
XX |||
FH Key 1..1434 Location/Qualifiers
FT CDS /tag=a
FT /product="Human peroxisome proliferator activated
FT receptor (PPAR) gamma1 protein"
XX
XX
PN US6294559-B1.
XX
PD 25-SEP-2001.
XX
PF 03-AUG-1998; 98US-0128142.
XX
PR 02-MAY-1996; 96US-016694P.
PR 18-APR-1997; 97US-0844007.
XX
XX (MER1 ) MERCK & CO INC.
XX
XX Smith RG;
PI
XX WPI: 2001-647265/74.
DR P-PSDB: AAE12867.
XX
XX Use of thiazolidinedione for treating cancer and viral infections -
XX
XX Example 3; Fig 1; 17pp; English.
XX
XX The invention relates to compounds and ligands that bind to human
XX peroxisome proliferator activated receptors (PPAR) gamma1 and gamma2.
XX The invention is useful for treating cancer and other disorders including
XX excessive cell proliferation and viral infection. The invention is also
XX directed to the use of PPAR gamma1 and gamma2 to identify compounds that
XX are antiproliferative, antiviral and antitumour agents. The invention
XX also relates to a method of treating cancer using a pharmaceutical
XX composition comprising thiazolidinedione in an amount sufficient to
XX modulate PPAR gamma1 and gamma2 activity. The present cDNA sequence
XX encodes human peroxisome proliferator activated receptor (PPAR) gamma1
XX protein related to the invention.
SQ
XX Sequence 1434 BP; 411 A; 343 C; 330 G; 350 T; 0 other:
XX
XX Alignment Scores:
XX Pred. No.: 2,42e-184 Length: 1434
XX Score: 2089.00 Matches: 415
XX Percent Similarity: 87.00% Conservative: 0
XX Best Local Similarity: 87.00% Mismatches: 0
XX Query Match: 59.55% Indels: 62
XX DB: 22 Gaps: 1
XX
XX US-09-931-007A-1 (1-688) x AAD21021 (1-1434)
QY 29 MetThrMetValAspThrGluMetProPheThrProThrAsnPheGlyIleSerSerVal 48
XX |||
DB 1 ATGACCATGTTGACACACAGATGATGCTGCGCCACCAACTTGGGATGAGTCGCG 60
QY 49 AspLeuSerValMetGluAspHisSerHisSerPheAspIleLysProPheThrVal 68
XX |||
DB 61 GATCTCTCCGTAATGGAACACACCTCCACTCTTGGATATCAACCCCTCACTACTGT 120
QY 69 AspPheSerSerIleSerThrProHisTyrGluAspIleProPheThrArGThrAspPro 88
XX |||
DB 121 GACTTCTCCAGCATTTCTCTCTCCACATTAGAGACATTTCCATTCCAAAGACATCA 180
QY 89 ValAlaAlaAspTyrIleTyrAspLeuLysLeuGlnGluTyrGlnSerAlaIleLysVal 108
XX |||
DB 181 GTGGTTGCAGATTACAGATGACCTGAACACTTCAAGAGATACCAAGTCAATCAAGTG 240
QY 109 GluProAlaSerProProTyrTyrSerGluLysThrGlnLeuTyrAsn----- 124
XX |||
DB 241 GAGCCTGCATCTCCACTTATTATTCTGAGAGACTCACTCTACAAATTAAGCCTCATGAA 300
XX
XX Human peroxisome proliferator activated receptor (PPAR) gamma1 cDNA.
XX
XX cytototoxic; antiproliferative; antiviral; cancer; cell proliferation;
KW
```

```
KW viral infection; pharmaceutical; thiazolidinedione; ss.
XX
XX OS Homo sapiens.
XX
XX FH Key 1..1434 Location/Qualifiers
XX CDS /tag=a
XX /product="Human peroxisome proliferator activated
XX receptor (PPAR) gamma1 protein"
XX
XX
XX US6294559-B1.
XX
XX 25-SEP-2001.
XX
XX 03-AUG-1998; 98US-0128142.
XX
XX 02-MAY-1996; 96US-016694P.
XX 18-APR-1997; 97US-0844007.
XX
XX (MER1 ) MERCK & CO INC.
XX
XX Smith RG;
PI
XX WPI: 2001-647265/74.
DR P-PSDB: AAE12867.
XX
XX Use of thiazolidinedione for treating cancer and viral infections -
XX
XX Example 3; Fig 1; 17pp; English.
XX
XX The invention relates to compounds and ligands that bind to human
XX peroxisome proliferator activated receptors (PPAR) gamma1 and gamma2.
XX The invention is useful for treating cancer and other disorders including
XX excessive cell proliferation and viral infection. The invention is also
XX directed to the use of PPAR gamma1 and gamma2 to identify compounds that
XX are antiproliferative, antiviral and antitumour agents. The invention
XX also relates to a method of treating cancer using a pharmaceutical
XX composition comprising thiazolidinedione in an amount sufficient to
XX modulate PPAR gamma1 and gamma2 activity. The present cDNA sequence
XX encodes human peroxisome proliferator activated receptor (PPAR) gamma1
XX protein related to the invention.
SQ
XX Sequence 1434 BP; 411 A; 343 C; 330 G; 350 T; 0 other:
XX
XX Alignment Scores:
XX Pred. No.: 2,42e-184 Length: 1434
XX Score: 2089.00 Matches: 415
XX Percent Similarity: 87.00% Conservative: 0
XX Best Local Similarity: 87.00% Mismatches: 0
XX Query Match: 59.55% Indels: 62
XX DB: 22 Gaps: 1
XX
XX US-09-931-007A-1 (1-688) x AAD21021 (1-1434)
QY 29 MetThrMetValAspThrGluMetProPheThrProThrAsnPheGlyIleSerSerVal 48
XX |||
DB 1 ATGACCATGTTGACACACAGATGATGCTGCGCCACCAACTTGGGATGAGTCGCG 60
QY 49 AspLeuSerValMetGluAspHisSerHisSerPheAspIleLysProPheThrVal 68
XX |||
DB 61 GATCTCTCCGTAATGGAACACACCTCCACTCTTGGATATCAACCCCTCACTACTGT 120
QY 69 AspPheSerSerIleSerThrProHisTyrGluAspIleProPheThrArGThrAspPro 88
XX |||
DB 121 GACTTCTCCAGCATTTCTCTCTCCACATTAGAGACATTTCCATTCCAAAGACATCA 180
QY 89 ValAlaAlaAspTyrIleTyrAspLeuLysLeuGlnGluTyrGlnSerAlaIleLysVal 108
XX |||
DB 181 GTGGTTGCAGATTACAGATGACCTGAACACTTCAAGAGATACCAAGTCAATCAAGTG 240
QY 109 GluProAlaSerProProTyrTyrSerGluLysThrGlnLeuTyrAsn----- 124
XX |||
DB 241 GAGCCTGCATCTCCACTTATTATTCTGAGAGACTCACTCTACAAATTAAGCCTCATGAA 300
XX
XX Human peroxisome proliferator activated receptor (PPAR) gamma1 cDNA.
XX
XX cytototoxic; antiproliferative; antiviral; cancer; cell proliferation;
KW
```

```

Qy 124 ----- 124
Db 301 GAGCCTTCAACTCCCTCATGGCAATGATGTGTCGTGGAGATAAAGCTTCTGGA 360
Qy 124 ----- 124
Db 361 TTTCACATGAGATTCAVCGTTGTGAAGATGCAAGGGTTTCTTCCGGAGAAACATTCAGA 420
Qy 125 ----- -Argasn 126
Db 421 TTGAAGCTTATCTATGACAGATGTGATCTTAACTGTCCGATCCACAAAAGTGAAT 480
Qy 127 LysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAlaIle 146
Db 481 AAATCTCAGTACTGTGCGTTTCAGAAATGCCCTTGACAGTGGGATCTCATATATCCATC 540
Qy 147 ArgPheGlyArgMetProGlnAlaGluLysGluLysLeuAlaGluIleSerSerAsp 166
Db 541 AGGTTTGGGCGGATCCACAGGCCGAGAAGGAGAGCTTTGGCGAGATCTCCAGTGAT 600
Qy 167 IleAspGlnLeuAsnProGlnSerAlaAspLeuArgAlaLeuAlaLysHisLeuTyrAsp 186
Db 601 ATCGACCAAGCTGAATCCAGAGTCCGCTGACCTCCGGCCCTGGCAAAACATTGTATATAC 660
Qy 187 SerTyrIleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLys 206
Db 661 TCATACATAAAGTCTCTCCGCTGACCAAGCAAGGCGAGGGCGATCTTGACAGAGAAAG 720
Qy 207 ThrThrAspLysSerPheProPheValIleTyrAspMetAsnSerLeuMetGlyLysAsp 226
Db 721 ACAACGAGCAAAATCACCATTCGTATCTATGACATGAATTCCTTATGATGGAGAAATAT 780
Qy 227 LysIleLysPheLysHisIleThrProLeuGlnGlnSerLysGluValAlaIleArg 246
Db 781 AAAATCAAGTTCAACACATCACCCCTGCGAGAGCAGACAAAGAGTGGCCATCCGC 840
Qy 247 IlePheGlnLysCysGlnPheArgSerValGluAlaValGlnGluIleThrGluTyrAla 266
Db 841 ATCTTTCAGGGCTGCGCAATTCGCTCCGCTGAGGCTGTGCAGAGATCACAGATATCC 900
Qy 267 LysSerIleProGlyPheValAsnLeuAsnLeuAsnAspGlnValThrLeuLeuLysTyr 286
Db 901 AAAACATTCCTGGTGTGTAAATCTGACTGAACGACCAAGTAACCTCTCTCAATATAT 960
Qy 287 GlyValHisGluIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyValLeu 306
Db 961 GGAGTCCACGAGATCATTTACACAATGCTGGCTCTTGATGAATAAAGATCGGTTCTC 1020
Qy 307 IleSerGlnGlnGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysProPhe 326
Db 1021 ATATCCGAGGGCCCAAGGCTTCATGCAAGGAGGATTTCTAAAGAGCTCGAAAGCTTTT 1080
Qy 327 GlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeuAsp 346
Db 1081 GGTGACTTATGAGACCCCAAGTTTGTGCTGCTGGAAGTTCAATGCACTGGAATTAGAT 1140
Qy 347 AspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGlyLeu 366
Db 1141 GACACCGACTGTGGCAATTTATTTATGCTGTCAATTTCTCAGGGAGACCGCCAGGTTTG 1200
Qy 367 LeuAsnValLysProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGluLeuGln 386
Db 1201 CTGATGTGGAAGCCCATTTGAGACATTCAGACAACTGCTNCAAGCCCTGAGAGCTCAG 1260
Qy 387 LeuLysLeuAsnHisProGlnSerGlnLeuPheAlaLysLeuLeuGlnLysMetThr 406
Db 1261 CTGAAGCTGAACCACTCGAGTCTCCACAGCGTTTGCCCAAGCTGCTCCAGAAATTCAGA 1320
Qy 407 AspLeuArgGlnIleValThrGluHisValGlnLeuLeuGlnValIleLysLysThrGlu 426
Db 1321 GACCTCAGCAAGATTGTCACGGAACACGTGCAAGCTACTGCAAGGTGATCAAGAAGCGAG 1380

```

```

Qy 427 ThrAspMetSerLeuHisProLeuLeuGlnGluIleTyrLysAspLeuTyr 443
Db 1381 ACAGACATGAGCTTCCACCCGCTCCTGCAGAGATCTACAGAGACTGTGATAC 1431
RESULT 11
AAK36522
ID AAK36522 standard; cDNA: 1811 BP.
AC AAK36522;
XX
XX 07-JUL-1999 (first entry)
DT
XX
XX Human PPAR-gamma1 coding sequence.
DE
XX
XX Nuclear receptor agonist; antagonist; identification; PPAR:
KW peroxisome proliferator activated receptor; ss.
XX
XX Homo sapiens.
OS
XX MO9918124-A1.
PN
XX
XX 15-APR-1999.
PD
XX
XX 06-OCT-1998; 98WO-US21049.
PE
XX
XX 07-OCT-1997; 97US-0061385.
PR
XX
XX (MERI ) MERCK & CO INC.
PA
XX
XX Cummings RT, Hermes JD, Moller DE, Zhou G;
PI WPI: 1999-263998/22.
XX
XX DR P-PSDB; AAY05471.
XX
XX PT Identifying nuclear receptor agonists and antagonists.
XX
XX PS Disclosure: Fig 9b; 60pp; English.
XX
XX CC This sequence encodes the human peroxisome proliferator activated
CC receptor-gamma1 (PPAR-gamma1).
CC
CC The invention relates to a method for identifying nuclear receptor
CC agonists and antagonists comprising measuring fluorescent resonance energy
CC transfer between fluorescent-labelled nuclear receptors and
CC co-activators. The method can be used for identifying agonists and
CC antagonist of nuclear receptors.
XX
XX SQ Sequence 1811 BP; 510 A; 433 C; 422 G; 446 T; 0 other.

Alignment Scores:
Pred. No.: 4,95e-182 Length: 1811
Score: 2065.50 Matches: 414
Percent Similarity: 85.71% Conservative: 6
Best Local Similarity: 84.49% Mismatches: 7
Query Match: 58.88% Indels: 63
DB: Gaps: 2

US-09-931-007a-1 (1-688) x AAK36522 (1-1811)
Qy 23 AlaAsnIleSerGlnGlnMetThrMetValAspThrGluMetProPheTrpProThrAsn 42
Db 155 GCCGCCGTGGCCCGCAGCAATGACCATGGTTGACACAGAGATCGCATCTTGGCCACCAAC 214
Qy 43 PheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPheAspIle 62
Db 215 TTGGGATCAGCTCCGTGATCTCTCGTAATGGAAGCACTCCCACTCTTGATATTC 274
Qy 63 LysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAspIlePro 82
Db 275 AAGCCCTTCATACGTGACTGTCTCCGCACTTCTACTCATTCAGAGACATTCACA 334
Qy 83 PheThrArgThrAspProValValAlaAspTyrLysTyrAspLeuLysLeuGlnGluTyr 102
Db 335 TTCAACAAGACAGATCCAGTGTGTCAGATTACAAAGTATGACCTGAAACTCAAGAGTAC 394

```

```

Oy 103 GlnSerAlaIleLysValGlnProAlaSerProProTyrTyrSerGlnLysThrGlnLeu 122
    |||
Db 395 CAAAGTGCATCAAGATGGAGCCTGCATCTCCACCTATTATTCTGAGAAGATCAGCTC 454
Oy 123 TyrAsn----- 124
    |||
Db 455 TACAAATTAAGCCTCAGTAAGAGCCCTCCAACTCCCTCATGGCAATTGAATGCTGCTGT 514
Oy 124 ----- 124
Db 515 GGAGATAAAGCTTTCGGATTTCACATGAGATTGATGCTTGGAAGATGCAAGGCTTC 574
Oy 124 ----- 124
Db 575 TTCGGAGAACATCAGATTGAAGCTTATCTATGACAGATGATGTTAAGTTCGGCATC 634
Oy 125 -----ArgAnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaGly 140
    |||
Db 635 CACAAAAAAGTAGAATAATATGTCTAGTACTGCTCGTTCCAGAAATGCCCTTGCGAGTGGG 694
Oy 141 MetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGlnLysGlnLysLeu 160
    |||
Db 695 ATGCTCATTAATGCCATCAGGTTTGCGGATCGCACAGCGCAGAGAGAGAGAGCTGTTG 754
Oy 161 AlaGlnLysSerSerAspIleAspGlnLeuAsnProGlnSerAlaAspLeuArg--Ala 179
    |||
Db 755 GCGAGATCTCCAGTATTCGACGACGATCCAGATCCGCTGACCTCCGTCAGGCGC 814
Oy 180 LeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLysAla 199
    |||
Db 815 CTGGCAAAACATTTGTACTCATCATTAAGTCTCTCCCTGCACCAACCAAGCGCG 874
Oy 200 ArgAlaIleLeuThrGlyLysThrAspLysSerProPheValIleTyrAspMetAsn 219
    |||
Db 875 AGGGCATTTTACAGAGAGAACACACAGACAAATCCATTCGTTATCTATCATCATGAAT 934
Oy 220 SerLeuMetMetGlyGlnAspLysIleLysPheLysHisIleThrProLeuGlnGln 239
    |||
Db 935 TCCTTAATGATGGAGAAATGAATCAATCAAGTCAAAACATCACCCTCGCAGAGCGAG 994
Oy 240 SerLysGlnValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGlnAlaVal 259
    |||
Db 995 ACCAAAGAGGTGGCCATCGCATCTTTCAGGGCTGCCAATTTCCCTCGTAGAGCTGTGG 1054
Oy 260 GlnGlnIleThrGlnTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 279
    |||
Db 1055 CAGGAGATCACAGAGATGCCAAAAGCATTCCTGCTTTGTAAATCTTGACTTGAAACGAC 1114
Oy 280 GlnValIleThrLeuLeuLysTyrGlyValHisGlnIleIleTyrThrMetLeuAlaSerLeu 299
    |||
Db 1115 CAAGTAACCTCTCTCAAAATATGAGAGTCCACGAGATCATTTACACATGCTGCGCTCTTG 1174
Oy 300 MetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGlnPheLeu 319
    |||
Db 1175 AAGAAATAAAGAGGGGTTCTCATATCCGAGGGCCCAAGGCTTATGACCAAGGAGAGTTTCTA 1234
Oy 320 LysSerLeuArgLysProPheGlnAspPheMetGlnProLysPheGlnPheAlaValLys 339
    |||
Db 1235 AAGAGCCTGCGAAAGCCTTTTGTGACTTTATGAGACCCCAAGTTGAGTTGCTGTGGAAG 1294
Oy 340 PheAsnAlaLeuGlnLeuAspSerAspSerAlaIlePheIleAlaValIleLeuLeu 359
    |||
Db 1295 TTCAATGCACCTGCAATTGATGACAGCGACTGGCAATATTATTGCTGCTATTCTTC 1354
Oy 360 SerGlyAspArgProGlyLeuLeuAsnValLysProIleGlnAspIleGlnAspAsnLeu 379
    |||
Db 1355 AGTGAAGACCGCCAGGTTTGTCTGAATGGAAGCCCATTTGAAGACATTCACAAACCACTTG 1414
Oy 380 LeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGlnLysSerGlnLeuPheAla 399
    |||
Db 1415 CTACAAGCCCTGGAGCTCAGCTGAAGCTGAACCAACCTGAGTCTTCACAGCTGTTTGGC 1474

```

```

Oy 400 LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValIleGlnHisValGlnLeu 419
    |||
Db 1475 AAGCTCTCCAGAAATGACAGACCTCCACAGATTTGTCACGGAACCTGCAGCTACTG 1534
Oy 420 GlnValIleLysLysThrLysThrAspMetSerLeuHisProLeuGlnGlnIleTyr 439
    |||
Db 1535 CAGGTGATCAAAAGACGAGACAGACATGAGTCTTCCACCCCTCTGACGAGATCTAC 1594
Oy 440 LysAspLeuTyrAlaIlePheAlaIleLeuThr 449
    |||
Db 1595 AAGGACTTGTACTAGACAGAGACTCCTGAGC 1624

```

RESULT 12

```

AAH76296
ID AAH76296 standard; cDNA; 1811 BP.
XX
AC AAH76296;
XX
XX 29-OCT-2001 (first entry)
XX
DE Human PPARgamma cDNA sequence.
XX
KW PAX8-PPARgamma1; oncogene; cytostatic; PAX8; PPARgamma1; cancer;
XX follicular carcinoma; PPARgamma; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 173..1609
    /*tag= a
XX
PN WO200152789-A2.
XX
PD 26-JUL-2001.
XX
PF 18-JAN-2001; 2001WO-US01664.
XX
PR 20-JAN-2000; 2000US-0177109.
XX
PR 14-AUG-2000; 2000US-0225079.
XX
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI Kroll TG, Fletcher JA;
XX
DR WPI: 2001-514487/56.
XX
DR P-PSDB: AAB85802.
XX

```

```

PT New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
    treating certain tumors or cancers, e.g. follicular carcinoma -
XX
XX Disclosure: Page 137-139; 145pp; English.
XX
XX

```

```

CC The invention relates to an oncogene designated PAX8-PPARgamma1 that
    contains a PAX8 coding region fused to PPARgamma1 coding region. The
    PAX8-PPARgamma1 polypeptides can be expressed by standard recombinant
    CC methodology. A PPARgamma1 ligand or agent is useful for treating a
    CC subject having a disorder characterized by the presence of a PAX8-
    CC PPARgamma1, where the disorder is cancer, e.g. follicular carcinoma.
    CC The PAX8-PPARgamma1 molecules are also useful for providing nucleotide
    CC and amino acid sequences useful for detecting the above disease. The
    CC present sequence represents a human PPARgamma1 polypeptide encoding cDNA.
XX
XX Sequence 1811 BP; 510 A; 433 C; 422 G; 446 T; 0 other;
XX

```

Alignment Scores:

```

Pred. No.: 4,95e-182 Length: 1811
Score: 2065.50 Matches: 414
Percent Similarity: 85.71% Conservative: 6
Best Local Similarity: 84.49% Mismatches: 7
Query Match: 58.88% Indels: 63
DB: 22 Gaps: 2

```

US-09-931-007A-1 (1-688) x AAH76296 (1-1811)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 25, 2003, 04:02:01 ; Search time 2471 Seconds

(without alignments)
4509.306 Million cell updates/sec

Title: US-09-931-007a-1

Perfect score: 3508

Sequence: 1 MGFELDSPIDPESDSFTDR.....KTEIDMSLHDLQIYKLDY 688

Scoring table:

BLOSUM62
Xgapop 10.0, Ygapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh

-Q=/cgn2_1/USPFO/US09931007/runatc_14022003_102553_4687/app.query.fasta.1.839

-DB=EST -QFMT=fastap -SUFFX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45

-DOCALLIGN=200 -THR_SCORE=pcct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09931007 @CGN.1.1698 @runatc_14022003_102553_4687 -NCPU=3

-NO_XLPHY -NO_MMAP -LARGEOUJURY -NEB_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120

-MARK.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -YGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estlin: *
4: em_estinu: *
5: em_estlov: *
6: em_estlpi: *
7: em_estro: *
8: em_hlc: *
9: gb_est1: *
10: gb_est2: *
11: gb_hlc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estum: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_iny: *
20: em_gss_pln: *
21: em_gss_vrt: *
22: em_gss_fun: *
23: em_gss_mam: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rtd: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1341	38.2	961	9	AL549707
2	1318	37.6	843	9	AL570116
3	1258	35.9	913	9	AL573383
4	1253	35.7	932	9	AL571005
5	1199	34.2	1026	9	AL523433
6	1198	34.2	793	12	BG742110
7	1136	32.4	826	10	BE535401
8	1122.5	32.0	900	9	AL549611
9	1098.5	31.3	1088	12	BG179310
10	1060.5	30.2	985	9	AL545177
11	996.5	28.4	944	14	BO883138
12	985	28.1	903	9	AL543579
13	985	28.1	872	14	BQ722874
14	933.5	26.6	901	13	BI820841
15	924.5	26.4	898	9	AL549037
16	919.5	26.2	574	9	AA625223
17	909.5	25.9	865	12	BG534958
18	905	25.8	551	12	BF192896
19	865	24.7	911	14	BQ930349
20	846	24.1	513	12	BF192893
21	837.5	23.9	959	9	AL523434
22	833	23.7	888	12	BG166774
23	813	23.2	719	14	BQ781472
24	812.5	23.2	932	14	BQ943112
25	798	22.7	694	14	BQ021210
26	767.5	21.9	1011	12	BE904444
27	757.5	21.6	986	14	BM923992
28	753.5	21.5	592	10	AM886988
29	752.5	21.5	700	9	AA673643
30	751.5	21.4	734	14	BO571382
31	740	21.1	450	13	BM433468
32	722	20.6	478	9	AI037566
33	716	20.4	533	10	AV605730
34	715	20.4	815	10	BE300340
35	699.5	19.9	938	10	BE300333
36	687	19.6	432	9	AA717831
37	684	19.5	746	9	AU131142
38	683	19.5	654	14	BM995901
39	682	19.4	965	14	BQ947830
40	681	19.4	430	12	BF522355
41	653	18.6	659	13	BI066593
42	649	18.5	823	13	BI524663
43	647	18.4	749	13	BI186429
44	646	18.4	524	10	AW911965
45	641.5	18.3	888	12	BG436329

ALIGNMENTS

RESULT 1
LOCUS AL549707/c 961 bp mRNA linear EST 16-FEB-2001
DEFINITION AL549707 LTL.NFL006.PL2 Homo sapiens cDNA clone CS0D1055Y11.3
ACCESSION AL549707
VERSION AL549707.1 GI:12865958
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 961)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization

JOURNAL COMMENT	Unpublished (2001) Contact: Genoscope
-----------------	--

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segreff@genoscope.cns.fr, Web : www.genoscope.cns.fr

FEATURES	Location/Qualifiers
source	1. .961

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="CS0D1055Y11"  
/clone_lib="LTI_NFL006_PL2"
```

	a	c	g	t	others
BASE COUNT	226	216	236	267	16
ORIGIN					

US-09-931-007A-1 (1-688) x AL549707 (1-961)

QY	163	ILSESRASPIIASEPGLINLEUASPROGLUSERLASPLEUALAGALALEUALYS	182
Db	953	GTCTCACTGATATGACCACTTAATCCAGAGTCGGCTACCTCGGGCCGTGGCAAAA	894
QY	183	HISLEUTYRASSETYTYRIELYSSEPHROLEUTHIRYSALATYSALATAGALALE	202
Db	893	CATTGTATATACCTACATCAATAAGTCTTCCCGCTGACCAAGMAAGGAGGGGATC	834
QY	203	LEUTHRGLYSTHR-THRASPLYSSEPROPIHEVALIETRYASPMETASERLEUKE	222
Db	833	TTTGACAGAAAGACAMACAGCAATACCATTCGTATATAGACATGAATTCCTTAAT	774
QY	222	TMEGLVGLUASPLYSILEYSPHELYSHISILETHPROLEUNGINDUSERLYSGI	242
Db	773	GATGGAGAAATATAATCAAGTTCAACACATCACCCCTCSAGSAGCAGAGCAAGA	714
QY	242	UVALALALEARGILEPHEGNGLYCYSGINPHEATRSERVALGUALAYALINGULI	262
Db	713	GGTGCCCATCCGCMATCTTTCAGGGCGTCCCGGTTCGCTGGAGAGCTTGCAGAGAT	654
QY	262	ETHGILTYTRIALYSSERIEPROGLYPHEVALASPLEUASPLEUASPCINVALAH	282
Db	653	CACAGACTATGCCAAAGCATCTCGGTTTGTGAATCTTGACTTGACACCAACAGTAAC	594
QY	282	RLLEULEUYSYRGLYALHISGLIUILEIETRYTHMETLEUALASERLEUETASNY	302
Db	593	TCTCCTCAATATGAGAGTCCACGAGATCATTTACACAACTCTGGCCTCGTTATGAATA	534
QY	302	SASPGLYVALLEUILESERGLUGLYNGLYPHEMETTHARGLUPHELEUYSSERIE	322
Db	533	AGATGGGGTCTCTATATCCGAGGGCCCAAGGCTTCATACACAGGAGAGTTCTTAAGAGCCT	474
QY	322	UARGLYSPROPHGLYASPMHEGTGUPPOLYSPHEGLUPHGLVALYSPHEASNL	342
Db	473	GCGAAAGCCTTTTGCTACTTTATATGAGGCCCAAGTTTGAATTGCTGTGAAGTTCATGC	414
QY	342	ALEUGLEUASPSASPERASPLEUALALEPHELEUALAVALLIETLEUSERGLYAS	362

Db	413	ACGGCAATTAGATGACAGCCACTTGGCAATATTATTATGCTCTATTATTCTCACTGAGCA	354
Oy	362	PATGProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeuLeuGlnAl	382
Db	353	CCGCCACAGTTTGGCTGGAATGTGAAGACCCATGGAAGRCATTCAGACACCACTGCTACACG	294
Oy	382	AlenGluLeuGlnLeuLysLeuAsnHisProGlnSerSerGlnLeuPheLAluLysLeu	402
Db	293	CCTGGAGCTCCMCGTGAAGCTGAACCAACCTGAGATCCTCAACAGCTGTTTGGCAAGCTGCT	234
Oy	402	UGLInLysMetThrAspLeuAArgLInIleValThrGlnHisValGlnLeuLeuGlnValIl	422
Db	233	CCAGAAATATKACMGCCCTCAGACAGATGTGTCMCGAACMCGTGMCTACTGCGAGGTGAT	174
Oy	422	elyLysLysThrGlnPArgAspMetSerLeuHisProLeuLeuGlnGlnLysLysAspLe	442
Db	173	CAGAGACAGCGAGACAGACATGAGTCTTCAACCCCTCTCGAGAGAGATCTACAAAGACTT	114
Oy	442	uTyrAlaTrpAlaIleLeuThr	449
Db	113	GTACTWGCAGAGACTCTGACG	92

RESULT 2	843 bp	mRNA	linear	EST 16-FEB-2001
AL570116/c				
LOCUS				
DEFINITION	AL570116	LT1_NFL006_P12	Homo sapiens cDNA clone CS0D10067110 3	
ACCESSION	AL570116	prime, mRNA sequence.		
VERSION	AL570116.1	GI:12926120		
KEYWORDS	EST.			
SOURCE	human.			

1 (bases 1 to 843)
 AUTHORS Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES	Location/Qualifiers
source	1. .843

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSODI006Y110"
/clone_lib="LTI_NFL006_PL2"
```

Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>

US-09-931-007A-1 (1-688) x AL570116 (1-843)

QY 177 LeuArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLys 196

```

|||||
Db      843 CTCGGCGCCCTGGCAAAACATTGTGTATGACATCATATAAAGCCTTCCCGCTGACCAAA 784
QY      197 AAlaValAlaArgAlaIleLeuThrArgIysThrThrAspLysSerProPheValIleYr 216
Db      783 GCAAAAGGCGAAGGGCGATCTTGACAGGAAAGACAAACAAATCCACATTCGTTATCTAT 724
QY      217 AspMetAsnSerLeuMetMetGlyGluAspLysIleLeuPheLysHisIleThrProLeu 236
Db      723 GACATTAATTCCTTATGATGGAGAGATTAATAATCAACTTCAACACATCACCCCTCG 664
QY      237 GluGluGlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerVal 256
Db      663 CAGGACCAAGCAAGAAAGAGTGGCCATCCCATCTTTCAGGGCTGCAGATTGCTCCGTG 604
QY      257 GluAlaValGlnGluIleThrGluYrAlaIysSerIleProGlyPheValAsnLeuAsp 276
Db      603 GAGGCTGTCCAGAGATCACAGATATGCCAAAGCATTCCTGGTTTGTAAATCTTGAC 544
QY      277 LeuAsnAspGlnValThrLeuLeuLysYrGlyValHisGlnIleIleYrThrMetLeu 296
Db      543 TTCAAGACCAAGTACTCTCTCTCAATATGAGATCCAGACATCATTTACACAAATGCTG 484
QY      297 AlaSerLeuMetAsnLysAspGlyValLeuIleSerGlyGlnGlyPheMetThrArg 316
Db      483 GCCTCTTGATGATTAAGATGGGGTTCATATCCGAGGCCAAGGCTTCATGACAAAG 424
QY      317 GluPheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPhe 336
Db      423 GAGTTTCTAAAGAGCTCGGAAAGCCTTTGTGTGACTTTATGAGAGCCAAAGTTTGAAGTTT 364
QY      337 AlaValLysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaVal 356
Db      363 GCGTGAATTCATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 304
QY      357 IleIleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGlnAspIleGln 376
Db      303 ATTATTCCTAGTGAAGACCGCCAGGTTGCTGAATGTGAAGCCCATGTAAGACATTTCAA 244
QY      377 AspAspLeuLeuGlnAlaLeuGlnLeuGlnIleLysLeuAsnHisProLysSerGln 396
Db      243 GAAACCTCTCTCAAGCCCTGGAGCTCCAGCTGAACTGAACCMCCCTAGCTCCTACAG 184
QY      397 LeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisVal 416
Db      183 CGTTTGGCAAGCTGCGCGAAGAAATGACAGACACTCAGACAGATTGTCCAGGAACGCG 124
QY      417 GlnLeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGln 436
Db      123 CAGCTACTCAGGTGATCAAGAAAGAGAGACAGACATGAGTCTTCCACCCGCTCCGAG 64
QY      437 GlnIleTyrlLysAspLeuYrAlaIleThrAlaIleLeuThr 449
Db      63 GAGATCTCAAGAGACTTGTACTAGACAGAGAGTCTGAGC 25

RESULT 3
AL573383/c 913 bp mRNA linear EST 16-FEB-2001
LOCUS      AL573383 LTI_NFL006.PL2 Homo sapiens cDNA clone CS001042YB22 3
DEFINITION AL573383 LTI_NFL006.PL2 Homo sapiens cDNA clone CS001042YB22 3
ACCESSION AL573383
VERSION    AL573383.1 GI:12932571
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 913)
AUTHORS   Li, W.-B., Gruber, C., Jesse, J. and Polayes, D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL    Unpublished (2001)
COMMENT    Genoscope - Centre National de Sequencage

```

```

FEATURES             BP 191 91006 EVRY cedex - France
                    Email: segret@genoscope.cns.fr, Web : www.genoscope.cns.fr.
                    Location/Qualifiers
                        source          1..913
                                         /organism="Homo sapiens"
                                         /db_xref="taxon:9606"
                                         /clone="CS001042YB22"
                                         /clone_1ib="LTI_NFL006.PL2"
                                         /tissue_type="placenta"
                                         /note="vector: pcnvsport 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pcnvsport 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : filiang@life.techn.com url :
http://fulllength.invitrogen.com"

BASE COUNT          222 a          206 c          215 g          268 t          2 others
ORIGIN

Alignment Scores:
Pred. No.:          1.33e-149          Length:          913
Score:              1258.00          Matches:          254
Percent Similarity: 98.46%          Conservative:      2
Best Local Similarity: 97.69%          Mismatches:       3
Query Match:        35.86%          Indels:           1
DB:                  9          Gaps:              0

US-09-931-007a-1 (1-688) x AL573383 (1-913)
QY      191 SerPheProLeuThrLysAlaValAlaIleLeuThrGlyLysThrThrAspLys 210
Db      912 TCCCTCCCGCTCAACCAAGGCGAGGCGATCTTACAGGAAAGCAACAGACAA 853
QY      211 SerProPheValIleTyrlAspMetAsnSerLeuMetGlyGluAspLysIleLysPhe 230
Db      852 TCACCATTCGTTATCATCATGATGATTCCTTATCATGATGGAGAAAGTAAATCAAGTTC 793
QY      231 LysHisIleThr-ProLeuGlnGlnSerLysGluValAlaIleArgIlePheGlnI 250
Db      792 AAACACATCACCCCCCTCCAGAGAGCAGCAAGAGCGGCATCCCATCTTCAGGG 733
QY      250 YcysGlnPheArgSerValGluAlaValGlnGlnIleThrGluYrAlaIysSerIlePr 270
Db      732 CTGCCAGTTTCCTCCGTGAGGCTGTGCAGAGATCACAGATGACCAAAAGCATTTCC 673
QY      270 cGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysYrGlyValHisGln 290
Db      672 TGGTTTGTAAATCTTGATTCAGACAGCAACTACTCTCCCAATATGAGATCCAGACA 613
QY      290 uIleIleYrThrMetLeuAlaSerLeuMetAsnLysAspGlyValLeuIleSerGlnI 310
Db      612 GATCATTTTACAAATGCTGCTCTTGTGATTAAGATGGGGTTCATATCCGAGGG 553
QY      310 yGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysProPheGlyAspPheMe 330
Db      552 CCAGAGCTTCAGACAAAGGAGTTTCTAAAGAGCCCTGCAAGAGCCTTTGGTGACTTTAT 493
QY      330 tGluProLysPheGluPheAlaValLysPheAsnAlaLeuGlnLeuLeuAspAspSerAspLe 350
Db      492 GGAGCCCAAGTTTGATTTGCTGTGATGATTCATGACATGACATGATGATGATGATGATG 433
QY      350 uAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGlyLeuLeuAsnValIy 370
Db      432 GGCATATATTATTTGCTGTCATATTCTCAGTGAAGACCCCGCAGGTTGCTGAATGTGAA 373
QY      370 sProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGlnGlnLeuLysLeuAs 390
Db      372 GCCCATGAGACATTTCAAGACACCTGCTACAAAGCCCTGGAGGCTCCAGCTTAACCTGAA 313
QY      390 nHisProLysSerGlnLeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArgGln 410

```

Db	312	CCACCCTAGTCTCTACAGCTCTTTGCCAGCTGCTCCAGAAATGACAGACCTTAGACA	253
Oy	410	mlleValThrGluHsValGlnLeuLeuGlnValIleLysLysThrGluThrAspMetSe	430
Db	252	GATGTCTACGGGAACAGCTGCACCTACTGTCAGGTGATCAAGAAGACGGAGACAGCATGAG	193
Oy	430	rlenuHsPProLeuLeuGlnGluIleTyrLysAspLeuYrAlaTrpAlaIleLeuThr	449
Db	192	TCTTACACCGCTCTCTGCAGAGAGATCTACAGGACATTGTACTACAGAGAGTCTTGAGC	135
RESULT 4	AL571005/c		
LOCUS	AL571005	932 bp	mRNA
DEFINITION	AL571005 LTI_NFL006_P12 Homo sapiens cDNA clone CS001028YJ24 3		linear EST 16-FEB-2001
ACCESSION	AL571005		
VERSION	AL571005.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 932)		
JOURNAL	Li, W.B., Gruber, C., Jesse, J. and Polayes, D.		
COMMENT	Full-length cDNA libraries and normalization		
	Unpublished (2001)		
	Contact: Genoscope		

FEATURES		SOURCE	
	Location/Qualifiers		
	1..932		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CSOD1028YJ24"		
	/clone_lib="TRI.NFL006.PL2"		
	/tissue_type="Placenta"		
	/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by life technologies, contact : Feng liang life technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com url : http://fulllength.invitrogen.com "		
BASE COUNT	230 a	207 c	210 g
ORIGIN		281 t	4 others

Alignment Scores:		
pred. No.:	6e-149	Length: 932
Score:	1253.00	Matches: 233
Percent Similarity:	97.70%	Conservative: 2
Best Local Similarity:	96.93%	Mismatches: 5
Query Match:	35.72%	Indels: 1
GB:	9	Gaps: 0

US-09-931-007A-1 (1-688) x AL571005 (1-932)

Qy	190	LYSSERPHPEPROLEUTHLlyslalalyslaalrghalaleleuthnglyLysThThThasp	209
Db	919	AAAKCCTTCGCGTGACCAAGCAAAAGCGAGGGGATCTTGACAGGAAACACACAGAC	860
Qy	210	LYSSERPROPHVALIILETYRASPmethsnbsrleumetGlyluAspLysIleLys	229
Db	859	AAATCACCATTTCGTATCTATGACATCAATTCCTTAAAGATGGGAGAAATCAATCAAG	800
Qy	230	PhelysHSLIETHPROLEUCLGLINLINSERLYSGluVALAlAlIlearyIIlepheln	249
Db	799	TTCAAAACATCACCCCTCGAGAGCAGAGCAAAAGAGGTGGCCATCCGATCTTTTCAG	740

QY	250	GIYSGINPhaIRSerSeraLGIuLaVaLGIuGLIeThrGLuTYrAlaLYSSerIle	269
Db	739	GGCTCGCAATTCCTGGCTGGAGGGTGGCGAGGATACAGAGATATCCAAAGCAAT	680
QY	270	ProGLyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysTyrGlyValHis	289
Db	679	CGTGTTTTGTAAATCTTACTTGAAGCAGCCAAAGTAATCTCTCCTCAAAATATGAGATCCAC	620
QY	290	GIuLIeIleTYrThMeLeuAlaSerLeuMeTasnLysAspGlyValLeuIleSerGlu	309
Db	619	GAGATCATTTTACACAAAGCTGGCTTCCTTGAGAAATAAAGATGGGTTCTCATATCCGAG	560
QY	310	GIYGLINGLYPheMetThrArgGluPheLeuLysSerLeuArgLysProPheGlyAspHe	329
Db	559	GGCCAAAGGCTTATGACAAAGGAGTTTCTAAAGAGCCTCGAAAGCCTTTGGTACTTT	500
QY	330	MetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeuAspAspSerAsp	349
Db	499	ATGGAGCCCAATTTGAGATTTCCTGTGAAAGTTCATGACTGGAATTTGATGACAGCGAC	440
QY	350	LeuAlaIlePheIleAlaValIleLeuSerGlyAspArgProGlyLeuLeuAsnVal	369
Db	439	TTGGCAATATTTATGTGCTGTCAATATTCATCGATGGAGACGCCAGGTTTGTGAATGTG	380
QY	370	LysProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGluLeuGlnLeuLysLeu	389
Db	379	AGGCCATTTGAAGACATTTCAAGCAACCTGCTACAAAGCCCTGGAGCTCCAGCTGAAGCTG	320
QY	390	AsnHisProGluSerSerGlnLeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArg	409
Db	319	AACCAACCCCTGAGTCTTACAGCTGTGTTGGCAAGCTGTCTCAAAAAATGACACACCTCACA	260
QY	410	GlnIleVal-TRGluHisValGlnLeuLeuGlnValIleLysLysThrGluThrAspHe	429
Db	259	CAGATTGTACACGGAACACGTCGACTACTGAGGTGATGACAAAGACGGAACAGACAT	200
QY	429	tSerLeuHisProLeuLeuGlnGluIleTyrLysAspLeuTyrAlaTAPAlaIleLeuTh	449
Db	199	GAGTCTTCACCCGCTCTCTGACAGAGATCATACAAAGCACTTGTCTAGCAGAGAGTCTTGAG	140
QY	449	r 449	
Db	139	c 139	
RESULT 5			
LOCUS	AL523433/c	1026 bp	mRNA linear EST 13-FEB-2001
DEFINITION	AL523433 LTL_NFL003.NBC3 Homo sapiens CDNA clone CS0DC004Y008 3		
ACCESSION	AL523433		
VERSION	AL523433.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLES	1 (bases 1 to 1026)		
JOURNAL	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.		
COMMENT	Full-length cDNA libraries and normalization		
	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRY cedex - France		
FEATURES	Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
source	Location/Qualifiers		
	1..1026		

Accession	Source	Organism	Accession	Source	Organism
204	ThrglyusthrtthAspIysSerProphetaValIletryrAspMetAsnSerIleuMetet	223	224	glyIuAspIysIleIasPheIysHisIlethrProIeugIngluInserIysgluVal	243
277	ACAGGAAGAAGACAAACAGCAAAATCACCATTCGTTACTCTATGACATGAATTCCTTAATGATG	336	337	GGAGAAAGATTAATAATCAAGTCAACACATCACACCCCTCGCAGAGCACAAGCAAAAGGTG	396
264	glyIuAlaIysSerIleProglIyheValasnuIeAspIeAsnAspGluValIthrIleu	283	397	GCACATCCCATCTTTCAGAGGCTGCAGATGTCCTCCGTGAGGAGCTGCAGAGATCACA	456
457	GAGTATGGCCAAAGCAATTCCTGGTTTGTAAATCTTGACTTGACGACCAAGTAACTCTC	516	284	IeulysTYrGlyValHisgluIleIletryrIleuIleAsnIleAsnIleuMetAsnIysAsp	303
517	CTCAATATMGAGGCTCCAGAGATCATTTACACAAATGCTGGCCTCTTGATGATTAATAT	576	304	glyValIleuIleSerGluGlyGlnGlyPheMetIhrArgIuPheIuIysSerIeuArg	323
577	GGGGTTCATATCCGAGGCGCAAGGCTTCATGACAAAGGAGTTCTTAAGAGCCTCGCA	636	324	IysProphetIyAspPheMetIuProIysPheglIupheIaValIyPheAsnIleu	343
637	AAGCCTTTGGGACTTTATGGAGCCCAAGTTGAGCTTGCTGTGAAGTTCATATGACATG	696	344	gluIeAspAspSerAspIeulalIe-PheIleIa-ValIleIleuSerGlyAspA	363
697	GAAATTAAGTGAAGGACTTGSCATATTTTATTTCTGTCTATTTCTCAGTGGAGAC	756	363	rgProglIyIeulAsnValIysProIeugIuAspIle	375
757	G-CCAGGTTTC-CTGATGTGAAGCNCATTTGAAGACATTT	792			
BE535401	BE535401.1	GI:9764046	826 bp	mRNA	linear
601058584F1	NIH-MGC_10	Homo sapiens	CDNA	clone	IMAGE:3445380 5',
BE535401	BE535401	EST			
human.					
Homo sapiens					
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
1 (bases 1 to 826)					
NIH-MGC	http://mgc.nci.nih.gov/				
National Institutes of Health, Mammalian Gene Collection (MGC)					
Unpublished (1999)					
Contact: Robert Strausberg, Ph.D.					
Email: cgapbs-re@mail.nih.gov					
Tissue Procurement: ATCC					
cDNA Library Preparation: Life Technologies, Inc.					
cDNA Library Arrayed by: Incyte Genomics, Inc.					
DNA Sequencing by: Incyte Genomics, Inc.					
Clone distribution: MGC clone distribution information can be					
found through the I.M.A.G.E. Consortium/BLN at:					
http://image.lnl.gov					
Plate: L1AM8415	row: k	column: 13			
High quality sequence stop: 36.					

FEATURES
source

```

"organism":"Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3445380"
/clone_lib="NTH.MGC_10"
/cell_line="MGC35"
/lab_host="DH10B"
note="Organ: cervix; Vector: pCMW-SPO6; Site:1: Nct1;
site.2: SalI; Cloned unidirectionally. Primer: oligo dT

```

Average insert size 1.5 kb. Library prepared by Life Technologies."			
BASE COUNT	229 a	187 c	221 g 189 t
ORIGIN			
Alignment Scores:			
Pred. No.:	4,42e-134	Length:	826
Score:	1136.00	Matches:	248
Percent Similarity:	94.66%	Conservative:	1
Best Local Similarity:	94.30%	Mismatches:	11
Query Match:	32.38%	Indels:	7
DB:	10	Gaps:	0

US-09-931-007A-1 (1-688) x BE535401 (1-826)

QY	129	glnTyrCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAlaIleArgPhe	149
Db	2	CAGTACTGTGGGTGCACAAATGCGCTTCAGTGGGAGTGTCTAATAGCCATCAGGTTT	61
QY	149	GlyArgMetProGlnAlaGlyLysGluLysLeuLeuAlaGluIleSerSerAspIleAsp	168
Db	62	GGGCGGATGCCACAGGCCGAGAGAGAAAGCTGTTGGCGGAGATCTCCAGTGAATATGCAC	122
QY	169	GlnLeuAsnProGluSerAlaAspLeuArgAlaLeuAlaLysHisLeuTyrAspSerTyr	186
Db	122	CAGTGAATCCAGATCCGCTGACTCTCGGGCCCTGGCAAAACATTTGTATGACTCATAC	181
QY	189	IleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLysThr	208
Db	182	ATAAAGTCCCTTCCCGCTACCAAAAGCGAGGGGATCTTGACAGGAAACACACA	243
QY	209	AspLysSerProPheValIleTyrAspMetAsnSerLeuMetGlyGluAspLysIle	228
Db	242	GACAATATCACCATTGCTATCTATCTAGTGAATTCCTTAATGATGGAGAAATAAATC	307
QY	229	LysPheLysHisIleThrProLeuGlnGlnLysIleSerLysGluAlaIleArgIlePhe	248
Db	302	AAGTCAAAACATCATCCCCCTGCAGAGACGAGAACAAAGAGTGGCCATCCGATCTTT	365
QY	249	GlnGlyCysGlnPheArgSerValGluAlaValGlnGluIleThrGluTyrAlaLysSer	266
Db	362	CAGGGCTCCCAAGTTTCGCTCCGTGGAGCTGTGCAGGATCACAGATATGCCAAAGC	423
QY	269	IleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysTyrGlyVal	288
Db	422	ATTCTGTG-TTTGTAAATCTTGACTTGACAGCACCAGTAACCTTCCTCAATATGGAGTC	487
QY	289	HisGluIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyValLeuIleSer	308
Db	481	CACGAGATCATTTACACAAATGCTGGCTCCCTTGTATGAATAAAGATGGGTTCTCATATCC	540
QY	309	GluGlyGln-GlyPheMetThrArgGluPheLeuLysSerLeuArgLysProPhe-GlyAla	322
Db	541	GAGGCCCAAGGGCTTCACAGACAGAGGAGCTTTCCTAAAGAGCTGGCAAAAGCCTTTTGGGTG	600
QY	328	sppPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeuAspAsp	348
Db	601	ACTTATATGAGCCCAAGTTTCAGTTTGGCTGTGAAGTTCAATGCACTGGAAATTAATGACA	666
QY	348	erAspLeuAlaIlePheIle-AlaValIleIleLeuSerGlyAspArgProGlyLeuLeu	367
Db	661	GGGACTTGGCAATATTTATGCGGGGCGCATTTTCTCAGGGGAGACCG-CCAGGTTCGTG	713
QY	368	AsnValLysProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGluLeuGlnLeu	387
Db	720	AATGTGAAGCCATT-GAAGACATTCAGGACACAACACTGTTCAAGCCG-GAGGCTACAGCTG	777
QY	388	Lys 388	
Db	778	ANG 780	
RESULT 8			
ALS49611			

LOCUS AL549611 900 bp mRNA linear EST 16-FEB-2001
 DEFINITION AL549611 LTR_NFL006.PL2 Homo sapiens cDNA clone CS0DI055Y11.5
 prime, mRNA sequence.
 ACCESSION AL549611
 VERSION AL549611.1 GI:12885763
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 900)
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.
 FEATURES
 source
 1..900
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DI055Y11"
 /tissue_type="placenta"
 /note="Vector: pCMVSPORT 6; site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by lite technologies. Contact : Feng Liang lite Technologies, a division of invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 266 a 201 c 214 g 218 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 2,72e-132 Length: 900
 Score: 1122.50 Matches: 229
 Percent Similarity: 77.78% Conservative: 2
 Best Local Similarity: 77.10% Mismatches: 3
 Query Match: 32.00% Indels: 63
 Gaps: 2

US-09-931-007a-1 (1-688) x AL549611 (1-900)

OY 94 LysTyRAspLeuLysLeuGlnGluTyrGlnSerAlaIleLysValGluProAlaSerPro 113
 |||||
 Db 3 AAGTATGACCTGAACCTCAAGAGTACCAAGTGCATCAAGTGGAGCTGCATCTCCA 62
 |||||
 OY 114 ProTyRysSerGluLysThrGlnLeuTyrAsn----- 124
 |||||
 Db 63 CCTATTATCTGGAAGACTGACTCAATATAGCCTCATGAAGACCTTCCAACTCC 122
 |||||
 OY 124 ----- 124
 Db 123 CTCATGCAATTGAATGCTGCTGTGGAGATAAAGCTTCTGCATTCTACTATGAGATT 182
 |||||
 OY 124 ----- 124
 Db 183 CATGCTGTGAAGATGCAAGGTTTCTCCGAGAACATCAGATTGAACCTTATCTAT 242
 |||||
 OY 125 -----ArgAsnLysCysGlnTyrCys 131
 |||||
 Db 243 GACAGATGTGATCTTAAGTGTGATCCACAAAAAAGTACAAATTAATGTGACACTGT 302
 |||||
 OY 132 ArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAlaIleArgPheGlyArgMet 151
 |||||
 Db 303 CGGTTTCAGAAATGCGCTTGCAGTGGGATGTCTCATATATGCCATCAGTTTGGCGGATG 362
 |||||
 OY 152 ProGlnAlaGluLysGluLysLeuAlaGluIleSerSerAspIleAspGlnLeuAsn 171

|||||
 Db 363 CCACAGGCCGAGAGAGAACTGCTGGCGAGATCTCCAGTATTCAGCACCTGAAT 422
 |||||
 OY 172 ProGlnSerAlaAspLeuArgAlaLeuAlaLysHisLysLeuTyrAspSerTyrIleLysSer 191
 |||||
 Db 423 CCAGATTCGCTGACCTCCGGCCCTGGCAAAACATTTGTATGACTCATATCAATAGTCC 482
 |||||
 OY 192 PheProLeuTrrLysAlaLysAlaArgAlaIleLeuTrrGlyLysThrThAspLysSer 211
 |||||
 Db 483 TTCCCGCTGACCAAGCAAGCGAGCGATCTTGCACGAAAGACACAGCAAAATCA 542
 |||||
 OY 212 ProPheValIleTyrAspMetAsnSerLeuMetGlyGluAspLysIleLysPheLys 231
 |||||
 Db 543 CCATTGCTTATCTATGACATGAATCTTATGATGAGGAAGATTAATCAAGTTCAA 602
 |||||
 OY 232 HistLeuTrrProLeuGlnGlnSerLysLysLysValAlaIleArgIlePheGlnGlyCys 251
 |||||
 Db 603 CACATCACCCCTGCAGACAGCAAGAGGTGGCCATCTCCATCTTTCAGCGCTGC 662
 |||||
 OY 252 GlnPheArgSerValGluAlaValGlnGluIleThrGluTrrAlaLysSerIleProGly 271
 |||||
 Db 663 CAGTTTCCCTCCGTGGAGAGCTGTGCAGAGATCACAGATGTGCCAAAGCATTCCTGCT 722
 |||||
 OY 272 PheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysTyrGlyValHisGluIle 291
 |||||
 Db 723 TTTGTAATCTTGACTTGAGACAGCAAGTACTCTCCCAATATGAGATCCACGAGATC 782
 |||||
 OY 292 IleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyValIleLysSerGlnGlyGln 311
 |||||
 Db 783 ATTTACACAAATGCTGGCTCTTGAATGAATGAAGTGGGTTCATATTCGAGAGGCCCA 842
 |||||
 OY 312 GlyPheMetTrrArgGluPheLeuLysSerLeuArgLysProPheGlyAsp 328
 |||||
 Db 843 GCGTTTCATGACAGAGGAGCTTCTAAGACT---GCGAAASCTTTTGGTGAC 890
 |||||

RESULT 9
 LOCUS BG179310 1088 bp mRNA linear EST 06-FEB-2001
 DEFINITION 602331030F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4432278 5',
 mRNA sequence.
 ACCESSION BG179310
 VERSION BG179310.1 GI:12686013
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1088)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: gspbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1AM10188 row: 1 column: 07
 High quality sequence stop: 658.
 FEATURES
 source
 1..1088
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4432278"
 /tissue_type="adipocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.4 kb. Library enriched for

full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

BASE COUNT 296 a 255 c 286 g 250 t 1 others

ALIGNMENT SCORES:

Pred. No.: 4,41e-129 Length: 1088
 Score: 1098.50 Matches: 242
 Percent Similarity: 96.47% Conservative: 4
 Best Local Similarity: 94.90% Mismatches: 8
 Query Match: 31.31% Indels: 6
 DB: 12 Gaps: 0

US-09-931-007A-1 (1-688) x BG179310 (1-1088)

OY 143 HISANALIAIARPHGHEGLYARQMETPROGLNALAGLULYSGLULYSLEULENALAGLU 162
 DB 3 CATNAATGCCATTCAGGTTTGCGGATGCCAGGCCGAGAGAGAGAGAGCTGTTGGCGGAG 62
 OY 163 ILESERSEASPILLEASPILEUNASPROGLUSERALIASPLEUARGLALEUNALALYS 182
 DB 63 AUCTCCAGATATGCACAGCGATCCAGATCCGCTGACCTCCGGGCCCTCGGCAANA 122
 OY 183 HISLEUTYRASPERTYRIILEYSSERPHEPROLEUTHLYSALALYSALARGALALIE 202
 DB 123 CATTTGTATGACTCATACATCAATAAGTCTCCGCTGACCAAGCAAGCGAGGCGGATC 182
 OY 203 LEUTHGLYLTHTHTHRASPLYSSEPROPHLEVALIILETYRASMETASNSERLEUMET 222
 DB 183 TTGACGAGGAAGCAACACAAATACCATTCGTATATGACATGGAATTCCTTAATG 242
 OY 223 METGLYLASPLYSILLEYSPHELYSHISILETHRPROLEUGLNUGLNSERTYGLU 242
 DB 243 ATGGGAGAGATNAATCAAGTTCAACACATCACCCCTCGAGAGAGACAGCAAGAG 302
 OY 243 VALALAIARGILEPHGLNGLCYSGINPHEARGSERVALGLUALVALGNGLUILE 262
 DB 303 GTGGCCATCCGATCTTTTCAGGGCTCCAGTTCCGCTCGTGAGAGCTGCGAGGATC 362
 OY 263 TTHTGLTYVALYSERILEPROGLYPHEVALASPLEUNASPSGLNVALIHTH 282
 DB 363 ACAGACTATGCCAAAGACATCTCGTGTGTGTAATCTTGACTGACACCAAGAACT 422
 OY 283 LEULEULSYFGLYVALHISGLUILEIERYTHMETLEUALASERLEUMELASLYS 302
 DB 423 CTCCTCAATATAGAGATGCCAGGATCATTTACAAATGCTGGCTCTGTATGATATAA 482
 OY 303 ASPEGLYVALLEULEISERGLUGLYNGLYPHEMETTHARGGLUPHEULEYSSERLEU 322
 DB 483 GATGGGTTCTCATATCCAGGGCCCAAGGCTTCATGACAAAGGAGTTTCTAAAGACCTG 542
 OY 323 ATGLYSPROPHGLYASPBHEMETGLUPOLYSPHEGLUPHEALVALYSPHEASNALA 342
 DB 543 CCAAGACCTTTGGTGACTTTATGAGGCCCAAGTTGAGTNGCTGTGAAGTTCAATGTA 602
 OY 343 LEUGLULENASPSPSERASPLEUALALIEPHEILEALVALIILELEUN-SERGLYAS 362
 DB 603 CTGGATTTAGATGACAG-CACTTGGCAATATTATATGCTGTGCTATATTTCTCCAGTGAGA 661
 OY 362 PARGLPROGLYLEULENASVALYSPROGLEULASPILEGLINASPSASPLEULEUNAL 382
 DB 662 CCGGGA-GGGTTGCTGAAAGTGAAGCCATTGAGACTT-CAAGACAACTGCT-AAAAC 718
 OY 382 ALEUGLULEUGLNULEULYSLEUNASHPISPROGLUSERSERGIN 396
 DB 719 TGTGAGAGTCCAGTTGAAGTTGAACA--CCTGAGTCTCTACAG 759

RESULT 10 985 bp mRNA linear EST 16-FEB-2001
 AL545177
 LOCUS AL545177 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D10281J24 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL545177

VERSION AL545177.1 GI:12877658
 EST.
 KEYWORDS human.
 SOURCE
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)

COMMENT

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. 985
 Location/Qualifiers

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1028YJ24"
 /clone_1lb="LTI_NFL006.PL2"
 /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifestech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 290 a 236 c 210 g 247 t 2 others

ALIGNMENT SCORES:

Pred. No.: 2.68e-124 Length: 985
 Score: 1060.50 Matches: 229
 Percent Similarity: 74.43% Conservative: 1
 Best Local Similarity: 74.11% Mismatches: 6
 Query Match: 30.23% Indels: 77
 DB: 9 Gaps: 2

US-09-931-007A-1 (1-688) x AL545177 (1-985)

OY 22 SERALIASNIIIESERGLN-----GLUMETHR 30
 DB :::::||||| ||| |||||||
 DB 59 ACCACAAATATACACAAAGCCATTTCACAAAGAGATMAAGCTTTACGAATGACC 118
 OY 31 METVALASPTHTGLUMETPROPHETRPROTHRASPNHEGLYIIESERSEVALASPLEU 50
 DB 119 ATGTTTGACACAGAGATGCCATTCTGGCCACCAACTTTGGGATCAGCTCCGTGATCTC 178
 OY 51 SERVALMETGLUASPHISERHISERIPHEASPILELYSPROPHETHTHTVALASPHE 70
 DB 179 TCCGTAATGAGAGACCACCTCCACTCTTTGTATATCAAGCCCTTACATACCTGTGACTTC 238
 OY 71 SERSERIIESERTHTPROHISTYRGLUSPILEPROPHETHTHARGTHASPROVALVAL 90
 DB 239 TCCACATTCTTACTCCACATTTACGAGACATTCATTCCACAGAACAGATCCAGTGT 298
 OY 91 ALASPTYRLYSTYRASPLEULYSLEUGLNUGLYRGINSERALALIELYSVALGLUPRO 110
 DB 299 GCAGATTACAAAGTATGACCTGAACTTCAAGAGTACCAAGTGCAATCAAGTGGAGCCT 358
 OY 111 ALASERPROPTOTYTYTSERGLULYSTRGLNULEUTYRASN----- 124
 DB 359 GCATCTCCACTTATATTATGTGAGAAAGCTCAGCTACAAATTAAGCTTCATGAAGAGCT 418
 OY 124 ----- 124
 DB 419 TCCAACTCCCTTCATGCAATTGAATGCTGTGCTGTGAGATAAAGCTTTCGATTTTCA 478
 OY 124 ----- 124

```
Db 479 TATGAGTTCATGCTTGTGAGAGATGCAAGGCTTCTCCGGAGAACATTCAGATTGAAG 538
Qy 125 -----ArgAnlyScys 128
Db 539 CTTATCTATGACAGATGATCTTAACCTGCGATCCACAAAAAAGTGAATATATGT 598
Qy 129 GlnTyrcysArphbcGlnLysCysLeuAlaValGlyMetSerHisAsnAlaIleArphe 148
Db 599 CAGTACTGTGGTTTCAGAAATGCTTGCAAGTGGGATGTCATATATGCAATCAGGTTT 658
Qy 149 GlyArgMetProGlnAlaGluLysGluLysLeuAlaGluIleSerSerAspIleasp 168
Db 659 GGGCGGATGCCACAGGCCAGAGAGAGAGCTGTGGCGGAGATCTCCAGTATATCGAC 718
Qy 169 GlnLeuAsnProGlnSerAlaAspLeuArgAlaLeuAlaLysHisIleuTyrcysSerTyrc 188
Db 719 CAGCTGAATCCAGAGTCCGCTTACCTCCGGGCCCTGGCAAAACATTTGTATGACTCATAC 778
Qy 189 IleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLysThrThr 208
Db 779 ATAAAGTCTCTCCGCTGACCAAGCAAGGCGAGGCGCATCTTGACAGAAAGACAACA 838
Qy 209 AspLysSerProPheValIleTyrcysMetAsnSerLeuMetMetGlyLysAspLysIle 228
Db 839 GACAAATCCACATTCGTATCTATGACATGAATTCCTTAATGATGGAGAAATATATC- 897
Qy 229 LysPheLysHisIleThrProLeuGlnGlnSerLysGluValAlaIleArgIlePhe 248
Db 898 AAGTTCAAACACATCACCCTCG-CAAGAGCAGAGCAAGCAAGAGTGCCCAT-CCCATCTTT 955
Qy 249 GlnGlyCysGlnPheArgSerValGlu 257
Db 956 CAGGCTGCAG-TTTCGCTCCGTGAG 981
RESULT 11
B0883138 944 bp mRNA linear EST 16-AUG-2002
LOCUS AGENCOURT_7982264 lupski_dorsal_root_ganglion Homo sapiens cDNA
DEFINITION B0883138
ACCESSION B0883138
VERSION B0883138.1 GI:22275146
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 944)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-rf@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Plate: LLM13578 row: j column: 11
High quality sequence stop: 707.
location/Qualifiers
1..944
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6186346"
/clone_id="lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:"
```

```
US-09-931-007a-1 (1-688) x B0883138 (1-944)
Qy 22 SerAlaAsnIleSerGln-----GluMetThr 30
Db 85 ACCACAAATATATCAACACAGCCATTTCCTCAACGAGAGTACGCTTTAAGCAAAATGACC 144
Qy 31 MetValAspThrGluMetProPheThrProThrAsnPheGlyIleSerSerValAspLeu 50
Db 145 ATGGTTGACACAGAGATGATGCATTTGGCCACCACTTTGGGATGAGTCCCTGGATCTC 204
Qy 51 SerValMetGluAspHisSerHisSerPheAspIleLysProPheThrValAspPhe 70
Db 205 TCCTGAATGAGAGACACACACCTCCCTTTGATATCAAGCCCTTGACATCTGATCTTC 264
Qy 71 SerSerIleSerThrProHisTyrcysLysAspIleProPheThrArgThrAspProValAla 90
Db 265 TCACACATTTCTACCTCCACATTCAGACAGCATTCATTCACAAAGATTCACAGTGGTT 324
Qy 91 AlaAspTyrcysTyrcysAspLeuLysLeuGlnGlnTyrcysValAlaIleLysValGluPro 110
Db 325 GCAGATTACAGATATGACCTGAACTTCACAGAGTACCAAAAGTGCATATCAAGTGGAGCTT 384
Qy 111 AlaSerProProTyrcysSerGluLysThrGlnLeuTyrcys----- 124
Db 385 GCATCTCCACCTTATTTATTTGAGAAGACTGACGCTCAATATCAAGCTCATGAAGAGCTT 444
Qy 124 ----- 124
Db 445 TCCAATCCCTCATGATGCAATTGATGCTGCTGTGTGAGATMAAGCTTGTGATTTCAC 504
Qy 124 ----- 124
Db 505 TATGAGTTCATGCTTTGTGAAGATGCAAGGCTTCTCCGAGACAATCAGATTGAAG 564
Qy 125 -----ArgAnlyScys 128
Db 565 CTTATCTATGACAGATGATCTTAACCTGCGATCCACAAAAAAGTGAATATATGT 624
Qy 129 GlnTyrcysArphbcGlnLysCysLeuAlaValGlyMetSerHisAsnAlaIleArphe 148
Db 625 CAGTACTGTGGTTTCAGAAATGCTTGCAAGTGGGATGTCATATATGCAATCAGGTTT 684
Qy 149 GlyArgMetProGlnAlaGluLysGluLysLeuAlaGluIleSerSerAspIleasp 168
Db 665 GGGCGGATGCCACAGGCCAGAGAGAGAGCTGTGGCGGAGATCTCCAGTATATCGAC 744
Qy 169 GlnLeuAsnProGlnSerAlaAspLeuArgAlaLeuAlaLysHisIleuTyrcysSerTyrc 188
Db 745 CAGCTGAATCCAGAGTCCGCTGACCTCCGGGCCCTGGCAAAACATTTGTATGACATCATAC 804
Qy 189 IleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLysThrThr 208
Db 805 ATAAAGTCTCTCCGCTGACCAAAAGGAGGAGGATCTTGACAGNAAGACACA 864
```

NOTE: Site-2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGCTCG-3' and 5'-GACTAGTCTGTGATGCGAGCGCGCCCTT(15)-3'. Size selected > 1 kb for average insert length 1.7 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

BASE COUNT 282 a 231 c 196 g 234 t 1 others

Alignment Scores:
Pred. No.: 3,78e-116 Length: 944
Score: 996.50 Matches: 204
Percent Similarity: 72.13% Conservative: 3
Best Local Similarity: 71.08% Mismatches: 74
Query Match: 28.41% Indels: 74
DB: 14 Gaps: 2

OY	209	AspLysSerProPheValIleTyrTrpAspMetAlaSerLeuMetGlyGluAspLysIle	228
Db	865	GACAAATC-CCATTGGTATCTATGATCATGAAATCCTTAAGATGGGAGAGATATAATC	923
OY	229	LysPheLysHisIleThrPro	235
Db	924	AMGTTCAACACATCACCACC	944
RESULT 12			
AL543579			
LOCUS	AL543579	903 bp	mRNA linear EST 16-FEB-2001
DEFINITION	AL543579 LTI_NFL006_PL2	Homo sapiens	CDNA clone CS01D1006Y110 5
ACCESSION	AL543579		prime, mRNA sequence.
VERSION	AL543579.1	GI:12876058	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 903)		
AUTHORS	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.		
TITLE	Full-length CDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 EVRI cedex - France		
	Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.		
FEATURES			
SOURCE	1..903		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS01D1006Y110"		
	/clone_lib="LTI_NFL006_PL2"		
	/tissue_type="Placenta"		
	/note="vector: pCMVSPORT 6; site:1: NotI; 1st strand cDNA		
	was primed with a NotI-oligo(dT) primer. Five prime end		
	enriched, double-stranded cDNA was digested with Not I and		
	cloned into the Not I and Eco RV sites of the pCMVSPORT 6		
	vector. Library was normalized. Library was constructed by		
	Life Technologies. Contact : Feng Liang Life Technologies,		
	a division of Invitrogen 9800 Medical Center Drive		
	Rockville, Maryland 20850, USA Fax : (1) 301 610 8371		
	Email : fliang@lifetech.com URL :		
	http://fulllength.invitrogen.com"		
BASE COUNT	249 a 228 c 220 g 205 t	1 others	
ORIGIN			
Alignment Scores:			
Prod. NO.:	7.71e-115	Length:	903
Score:	986.00	Matches:	200
Percent Similarity:	75.66%	Conservative:	2
Best Local Similarity:	74.91%	Mismatches:	3
Query Match:	28.11%	Indels:	63
		Gaps:	1
US-09-931-0074-1 (1-688) x	AL543579 (1-903)		
OY	23	AlaAsnIleSerGlnGluMetThrMetValAspTrnGluMetProPheTrpProThrAsn	42
Db	97	GCCGCGGCGCGCGGAGAAATGACCATGGTTGGACACAGAGATGCCATCTGGCCACCAAC	156
OY	43	PheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPheAspIle	62
Db	157	TTTGGGATCAGCTCCGATGATCTCCGTAAGGAAGACCACTCCCACTTGTGATATC	216
OY	63	LysProPheTrnThrValAspPheSerSerIleSerTrnProHisTyrGlnAspIlePro	82
Db	217	AAGCCCTTACACTACGTTGACTTCTCCAGACATTTCTACTCCACATTAGCAAGACATTTCA	276
OY	83	PheThrArgTrnAspProValValAlaAspTyrLysTyrAspLeuLysLeuGlnIuTyr	102

Db	277	TTCCACAGACAGATCCACTGGTTGCAGATTTACAAAGTATGACCTGAACTTCAAGAGTAC	336
QY	103	gInserAlaIleIyValJcIuPrAlaSerProProTyTyrSerGluYshrgInIeu	122
Db	337	CAAAATGCANTCAAAAGTGGAGCCATGCATCTCCACCTTATTATTGTGAAGAAGCTCAGCTC	396
QY	123	TYrAsn-----	124
Db	397	TTCATTAAGCCTCATGAGAGCCCTTCCAACTCCCTCATGCGCAATTGAATGTGTGTCTGT	456
QY	124	-----	124
Db	457	GGAGATAAAGCTTCTGGATTTCATATGAGATTTCATGCTTGTGAAGGATGCAAGGTTTC	516
QY	124	-----	124
Db	517	TTCCGGAGACACATCAGATTGAAGCTTATCTATGACAGATGTGATCTTAACTGTCCGATC	576
QY	125	-----ArgAnIyScySglInTyCYCYArGPhgInIyScySLeuAlaValGIy	140
Db	577	CACAAAAAAGTAGAAATTAATGTGATGCTGCTGTGCGGTTTCACAAATAGCTTGACAGTGGG	636
QY	141	MetSerHisAsnAlaIleArgPheGIyArGmetProGInAlaGIuIySGLuYsLeuIeu	160
Db	637	ATGTCATCAAAAGCCATCAGCTTGTGGCGCGAGTGA-CAGGCCAGAGAAGAGAGCTGTG	695
QY	161	AlaGIuIleSerSerAspIleAspIleAsnProGInSerAlaAspLeuArgAlaIeu	180
Db	696	GGGAGATCTCCAGTGAATATGCACCAAGCTGATCCAGAGTCCGCTGACCTCCGGCCCTG	755
QY	181	AlaIySHisLeuTyRAspSerTyRleIySserPheProLeuThrLysAlaIySAlaArg	200
Db	756	GCAAAACATTTGTATGACATACATCAATAAAGTCTCTCCGCTGACCAAGCAAGCCGAGG	815
QY	201	AlaIleuThrGIySThrThraSplySerProPheValIleTyRAspMetAsnSer	220
Db	816	GGGATCTTGACAGGAAGACACAGACACATCAACATTCGATTATGATGACATGAATTC	875
QY	221	LeuMetMetGIyGIuAspIyS	227
Db	876	TTAATGATGGAGAGAAATAAA	896
RESULT 13			
LOCUS	BO722874	872 bp	mRNA
DEFINITION	AGENCOURT_8099880	Lupski_sympathetic-trunk	Homo sapiens cDNA clone
ACCESSION	BO722874	IMAGE:6190020	5', mRNA sequence.
VERSION	BO722874.1	GI:21861771	EST.
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE	1 (bases 1 to 872)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/.		
COMMENT	Unpublished (1999)		
	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cga@bbs.femail.nih.gov		
	Tissue Procurement: Dr. James R. Lupski		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	Plate: LRAM13588	row: c	column: 13
	High quality sequence stop: 601.		
FEATURES			
Source	Location/Qualifiers		
	1..872		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		

```

/clone="IMAGE:6190020"
/clone_lib="lupski_sympathetic_trunk"
/sex="male"
/tissue_type="sympathetic trunk"
/dev_stage="adult, 16 yr"
/lab_host="DH10B"
/Note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGAGCCGCGGCG-3' and
5'-GACATGTTCTAGATCCGACGCGCCGCTT(15)-3'. Size selected >
1 kb for average insert length 1.9 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine); available through Life
Technologies."

```

BASE COUNT 245 a 195 c 221 g 211 t

ORIGIN

Alignment Scores:

Pred. No.:	9.74e-115	Length:	872
Score:	985.00	Matches:	202
Percent Similarity:	95.77%	Conservative:	2
Best Local Similarity:	94.84%	Mismatches:	5
Query Match:	28.08%	Indels:	4
DB:	14	Gaps:	2

US-09-931-007a-1 (1-688) x BQ722874 (1-872)

```

QY 124 AsnAtgAnlnYsCysgLnTyrCysArghPheGlnYsCysLeuAlaValGlyMetSerHis 143
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 189 ACTAGAAATAAATGTCAGACTGTGCTTCAGAAATGCTTCGAGTGGGATGTCAT 248

QY 144 AsnAlaIleArPhgLYArGmetProGlnAlaGlnYsGlnYsLeuLeuAlaGlnUlle 163
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 249 AATGCCATAGGTTGGGGGATGCCACAGCGCGAAGGAAAGAACTGTTGGCGGAGATC 308

QY 164 SerSerAspIleAspGlnLeuAnProGlnSerAlaAspLeuArGAlaLeuAlaYsHis 183
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 309 TCCAGGATATGACACAGTCGATCCAGATCCGCTGACCTCGGCGCTGGCAAAACAT 368

QY 184 LeuTYrAspSerTYrIleYsSerPheProLeuThrLYsAlaLYsAlaArGAlaIleLeu 203
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 369 TTGTATGACTCATATAAAGTCTTCGCCGTGACCAAGCAAGCGAGGCGGATCTTG 428

QY 204 ThrGlyLysThrThrasPlySerProPheValIleTYrAspMetAsnSerLeuMet 223
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 429 ACAGAAAGACACAGACAAATCACCATTCGTTATCTATGACATGAATTCCTTAATGATG 488

QY 224 GlyGlnAspLYsIleLYsPheLYsHisIleThrProLeuGlnGlnSerLYsGlnVal 243
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 489 GGAGAAGATAAATCAAGTCAACACATCACCCCTCGACAGACAGACAGAAAGAGTG 548

QY 244 AlaIleArghIlePheGlnGlyCysGlnPheArGSerValGlnAlaValGlnUlleThr 263
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 549 GCCATCCGATCTTCAGGGCGCCAGATTCCCTCCGTGAGAGCGTGTGAGAGATGACA 608

QY 264 GlnTYrAlaLYsSerIleProGlyPheValAsnLeuAspLeuAspGlnValThrLeu 283
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 609 GAGTATGCCAAAGCATTCCTGCTGTTTGTGAATCTTGATGAACACCAATTAATCTGC 668

QY 284 LeuLYsTYrGlyValHisGlnUlleIleTYrThrMetLeuAlaSerLeuMetAsnLYsAsp 303
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 669 CTCATATATGAGGATCGACAGATCATTTACACATGCTGGCTCTTGTATGATGAATGAT 728

QY 304 GlyValLeuIleSerGlnUlyGlnGlyPheMetThrasGlnUleLeuLYs---SerL 322
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 729 GGGGTTCTCATATCCGAGGGCGCAGGCTTCATGAGCAAGGAGTTTCTAAAGAGCCT 788

QY 322 euArgLYsProPhe---GlyAspPheMetGluPro 332
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 789 GCGGAAAGCCCTTTTGTGGGACTTTTATGGACCC 823

```

RESULT 14
B1820841 901 bp mRNA linear EST 04-OCT-2001
LOCUS 603034063p1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:517551 5',
DEFINITION mRNA sequence.
ACCESSION B1820841
VERSION B1820841.1 GI:15932391
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 901)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNALS Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
Plate: LRAM11436 row: m column: 16
High quality sequence stop: 876.

FEATURES

Source

1. 901
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5175351"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis. Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH-MGC Library."

BASE COUNT 255 a 214 c 189 g 243 t

ORIGIN

Alignment Scores:

Pred. No.:	3.95e-108	Length:	901
Score:	933.50 <td>Matches:</td> <td>188 </td>	Matches:	188
Percent Similarity:	73.26% <td>Conservative:</td> <td>1 </td>	Conservative:	1
Best Local Similarity:	72.87% <td>Mismatches:</td> <td>6 </td>	Mismatches:	6
Query Match:	26.61% <td>Indels:</td> <td>63 </td>	Indels:	63
DB:	13 <td>Gaps:</td> <td>2 </td>	Gaps:	2

US-09-931-007a-1 (1-688) x B1820841 (1-901)

```

QY 1 MetGlyGlnThrLeuGlyAspSerProIleAspProGlnSerAspSerPheThrAspThr 20
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 127 ATGGGTGAACCTGGGAGATTCTCTAATGACCCAGAAACGATTCCTTCACTGATCA 186

QY 21 LeuSerAlaAsnIleSerGlnGlnUleThrMetValAspThrGlnUleProPheThrPro 40
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 187 CTGTGTGAACATATATCAAGAAATGACATGCTGTGACACAGATGSCATTCGCGCC 246

QY 41 ThrAsnPhgIlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 247 ACCAATCTTTGGAGATCAGCTCCGTGATCTCCGTAATAGGAAGACATCCCATCTTT 306

QY 61 AspIleLYsProPheThrThrasPheSerSerIleSerThrProHisTYsGlnAsp 80
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB 307 GATATCAAGCCCTTCTACACTGTGACTTCTCCAGCATTTCTACTCCACATTAAGAGAC 366

```

```
QY 81 lIePheThrArqThrasProValValAlaAspTyrLysTyrAspLeuLysLeuGln 100
Db ATTCATTCACAGAAAGACAGATCCAGTGGTTCAGATTACAGATATACCTGAAACTTCAA 426
QY 101 GluTyrGlnSerAlaIleLysValGluProAlaSerProGlyTyrSerGluLysThr 120
Db GAGTACCAAGAGTCACATCAAAAGTGAGCCTGCATCCACCTTATATTCTGAGAGACT 486
QY 121 GlnLeuTyrAsn----- 124
Db CAGCTCTACATATAGCCTCATGAAGACCTTCCAACTCCCTCATGSCAATGAATGTCTG 546
QY 124 ----- 124
Db GTCGTGAGATTAAGCTTCTGGATTTCACATAGATTCATCTGTGAGAGATGCAAG 606
QY 124 ----- 124
Db GGTTCCTCCGAGAACATCAGATTGAAGCTTATCTATGACAGATGTGATCTTAACGT 666
QY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
Db CGGATCCACAAAAAGTAGAATAATATGTCAGTACTGTCGTTTCAGAAATGCTTGA 726
QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158
Db GGGGAGATGTCATATATCCATCAGTGGGCGGATGCCACAGCCGAGAGAGAGAAAG 786
QY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeu---AsnProGlnSerAlaAspLeu 177
Db CTGTTGGCGAGATCTCCAGTGTATCGACACAGCGTGAATCCCGAGAGTCCGCTGACCTC 846
QY 178 ArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThr 195
Db CGGGGCCCTGCGAAAAACATTGTATGACTCATATCAATGAAGTCCGCCGCTGACA 900
RESULT 15
AL549037 898 bp mRNA linear EST 16-FEB-2001
LOCUS AL549037 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1042YB22 5
DEFINITION AL549037 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1042YB22 5
ACCESSION AL549037
VERSION AL549037.1 GI:12884632
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 898)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..898
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1042YB22"
/clone_lib="LTI_NFL006.PL2"
/ribose_type="Placenta"
/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
```

```
BASE COUNT 269 a 211 c 177 g 228 t 13 others
ORIGIN
Alignment Scores:
Pred. No.: 5,56e-107 Length: 898
Score: 924.50 Matches: 190
Percent Similarity: 68.95% Conservative: 1
Best Local Similarity: 68.59% Mismatches: 13
Query Match: 26.35% Indels: 73
DB: 9 Gaps: 2
US-09-931-007a-1 (1-688) x AL549037 (1-898)
QY 22 SerAlaAsnIleSerGln-----GluMetThr 30
Db ACCACAAATATACACACAGCCATTTTCACAAACGAGTCAGCTTATACCAATATGAC 117
QY 31 MetValAspThrGluMetProPheThrProThrAsnPheGlyIleSerSerValAspLeu 50
Db ATGGTTGACACAGACAGATGCCATTCGGCCACCACTTGGGATCAGCTCCGATGATC 177
QY 51 SerValMetGluAspHisSerHisSerPheAspIleLysProPheThrValAspPhe 70
Db TCGATATGGAAGACCACTCCACTCTTTGATATCAAGCCTTCATCACTACTGTGACTTC 237
QY 71 SerSerIleSerThrProHisTyrGluAspIleProPheThrArgThrAspProVal 90
Db TCCAGCATTTTACTATCCACATTCAGAGACATTCATTCACAAAGAACATCAGTGTT 297
QY 91 AlaAspTyrLysTyrAspLeuLysLeuGlnGluTyrGlnSerAlaIleLysValGluPro 110
Db GCGATTTACAAATATGACCTGAAACTTCAAGAGTACCAAGTGCAATCAAAAGTGAGCCT 357
QY 111 AlaSerProProTyrTyrSerGluLysThrGlnLeuTyrAsn----- 124
Db GCATCTCCACCTTATTATCTGAGAAAGACTCAGCTTCAATATACCTCATGACAGCT 417
QY 124 ----- 124
Db TCCAACTCCCTCATGAGCAATTAATGTGCTGTGTGAGATAAGCTTCTGATTTTCA 477
QY 124 ----- 124
QY 478 TATGAGTTCAATCTTGTGAGAGATGCAAGGGTTCTCCGGAGAACATCAGATTGAAG 537
QY 125 -----ArgAsnLysCys 128
Db CTATATCTATGACAGATGTATCTTAACCTCGGATCCACAAAAAGTAGAATAATAT 597
QY 129 GlnTyrCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAlaIleArgPhe 148
Db CAGTACTGTCGGTTTCASAAATSCCTTGGCAGGGGATGTCTCATATGCGCATGAGTTT 657
QY 149 GlyArgMetProGlnAlaGluLysGluLysLeuLeuAlaGluIleSerSerAspIleAsp 168
Db GGGCGGATGCCACAGCCGAGAGAGAGCGTTTSGCGAATATCCAGISATATTCAC 717
QY 169 GlnLeuAsnProGlnSerAlaAspLeuArgAlaLeuAlaLysHisLeuTyrAspSerTyr 188
Db CAGCTGAATCCAGAGATCCCGTGAACCTCCGGGCCCTGGCAAAAACATTTGTATGACCT 777
QY 189 IleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLysThrThr 208
Db ATTAAGTCTCTCCGCTGCCACCAAGCAGGCGGATCTTGACASGAAGAGAMAMA 837
QY 209 AspLysSerProPheValIleTyrAspMetAsnSerLeuMetMetGlyGlu 225
Db GACAATATCACCAATGCTTAATATGACATGAATTCCTTAATGATGGAGAG 888
Search completed: February 25, 2003, 06:13:34
Job time : 2484 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 25, 2003, 04:02:41 ; Search time 273 Seconds
(without alignments)
1415.414 Million cell updates/sec

Title: US-09-931-007A-1
Perfect score: 3508
Sequence: 1 MCELTGSDSPIDESPDSFTDT.....KTEFTDMSLHPLGLQIYKDLV 688

Scoring table:
BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 442118 seqs, 280819700 residues
Total number of hits satisfying chosen parameters: 884236

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=framed_p2n_model -DBY=xlh
-Q=/cgn2_1/USPFO.spool/US0931007/runat_14022003_102554_4718/app_query.fasta.1.839
-DB=Published.Applications_NA -OFMT=fastap -SUFFIX=p2n.inpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62
-TRANS=human40.cdi -LIST=45 -DCCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US0931007.ecgn.1.1.41 @runat_14022003_102554_4718
-NCPU=6 -ICPU=3 -NO_XLIPY -NO_MMAR -LARGOQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	2231	63.6	1608	US-09-765-111A-15
2	2196.5	62.6	1679	US-10-109-886-5
3	2106.5	60.0	2596	US-09-765-111A-22
4	2106.5	60.0	2625	US-09-765-111A-5

5	2099	59.8	2334	US-09-765-111A-1	Sequence 1, Appl1
6	2094	59.7	2523	US-09-765-111A-3	Sequence 3, Appl1
7	2065.5	58.9	1811	US-09-765-111A-26	Sequence 26, Appl1
8	2065.5	58.9	1811	US-09-880-107-2212	Sequence 2212, Ap
9	2065.5	58.9	2260	US-09-816-828-8	Sequence 8, Appl1
c 10	2065.5	58.9	2307	US-10-125-237-5	Sequence 5, Appl1
c 11	2065.5	58.9	2329	US-09-816-828-9	Sequence 9, Appl1
12	2065	58.9	2005	US-10-142-373-1	Sequence 1, Appl1
13	1101.5	31.4	3301	US-09-954-456-554	Sequence 554, App
14	1076.5	30.7	1323	US-10-013-807-1	Sequence 1, Appl1
15	702	20.0	428	US-10-033-528-1855	Sequence 1855, Ap
16	562	16.0	415	US-09-925-299-256	Sequence 256, Appl
17	512.5	14.6	611	US-09-925-299-109	Sequence 109, App
18	472	13.5	451	US-09-864-761-19039	Sequence 19039, A
19	406.5	11.6	3628	US-09-917-800A-1454	Sequence 1454, Ap
20	398.5	11.4	343	US-10-046-935-990	Sequence 990, App
21	398.5	11.4	343	US-09-878-178-990	Sequence 990, App
22	392	11.2	6282	US-10-108-605-212	Sequence 212, App
23	385.5	11.0	5970	US-10-108-605-210	Sequence 210, App
24	384	10.9	2880	US-10-044-090-361	Sequence 361, App
25	345	9.8	458	US-09-864-761-2304	Sequence 2304, Ap
26	315	9.0	2907	US-09-954-456-524	Sequence 524, App
27	313	8.9	1577	US-09-797-727-1	Sequence 1, Appl1
28	307	8.8	1679	US-10-044-090-316	Sequence 316, App
29	305	8.7	1528	US-09-962-832-109	Sequence 109, App
30	305	8.7	1528	US-09-880-107-3328	Sequence 3328, Ap
31	300	8.6	2066	US-09-895-840-1	Sequence 1, Appl1
32	281.5	8.0	1431	US-10-188-721-2	Sequence 2, Appl1
33	278	7.9	2218	US-09-880-107-3411	Sequence 3411, Ap
34	276.5	7.9	1934	US-09-909-446-1	Sequence 1, Appl1
35	276.5	7.9	1934	US-09-909-325-1	Sequence 1, Appl1
36	276.5	7.9	1934	US-09-909-326-1	Sequence 1, Appl1
37	274.5	7.8	1841	US-10-013-823-1	Sequence 1, Appl1
38	271.5	7.7	1594	US-09-964-824-559	Sequence 559, App
39	266	7.6	1752	US-09-965-703-7	Sequence 7, Appl1
40	266	7.6	1878	US-09-965-703-6	Sequence 6, Appl1
41	266	7.6	1878	US-09-965-703-6	Sequence 6, Appl1
42	266	7.6	2241	US-09-042-488B-4	Sequence 4, Appl1
43	266	7.6	2241	US-09-042-488B-6	Sequence 6, Appl1
44	266	7.6	3126	US-09-042-488B-8	Sequence 8, Appl1
45	265	7.6	1404	US-09-965-703-62	Sequence 62, Appl1

ALIGNMENTS

RESULT 1
US-09-765-111A-15
; Sequence 15, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroil, Todd G.
; TITLE OF INVENTION: PAX8-PARAGAMMA NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAR
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)...(1608)
; US-09-765-111A-15

Alignment Scores:

DD 391 CAGAGCAGAGAGG1GGCCATCCGCACTTCAGGGCTGGCCAGTTTCGCTCCG1GGAGGCT 103

Alignment Scores:

Score:	2196.50	Matches:	439
--------	---------	----------	-----

Percent Similarity: 87.15% Conservative: 2
 Best Local Similarity: 86.76% Mismatches: 2
 Query Match: 62.61% Indels: 63
 DB: 12 Gaps: 2

US-09-931-007a-1 (1-688) x US-10-109-886-5 (1-1679)

```

OY 1 MetGlg1ThrLeuGlyAspSerProIleAspProGluSerAspSerPheThr 20
    |||||||
Db 159 ATGGGGAACCTGGGAGATTCTCTATTGACCCAGAAAGCGATTCTTCAGTATCA 218
OY 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetPro 40
    |||||||
Db 219 CTGTGCAACATATCACAGAAATGACCATGTTGACACAGAGATTCACATTCGCCC 278
OY 41 ThrAspPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerPhe 60
    |||||||
Db 279 ACCAACTTGGGATCGAGCTCGGTGATCTCCGTAAAGAACCCACCTCTCTT 338
OY 61 AspIleLysProPheThrThrValAspPheSerIleSerThrProHisThrGluAsp 80
    |||||||
Db 339 GATATCAAGCCCTTCACTACTGTGTGACTTCTCCAGCATTTCTACATTCAGAGAC 398
OY 81 IleProPheThrArgThrAspProValValAlaAspTyrLysTyrAspLeuLysGln 100
    |||||||
Db 399 ATTCATTCACAAGACAGATCCAGTGTTCAGATTACAGATGACCTGAAACTTCAA 458
OY 101 GluTyrGlnSerAlaIleLysValGluProIleAspProTyrTyrSerGluLysThr 120
    |||||||
Db 459 GAATACCAAGTGCATCAAAAGTGAGCCTGCATCCACCTTATTATTCGTGAGAGACT 518
OY 121 GlnLeuTyrAsn----- 124
    |||||||
Db 519 CAGCTTACAAATAGCCTCATGAGAGCCTTCCAACTCCCTCATGSCAATTGATGTCT 578
OY 124 ----- 124
Db 579 GTCGTGAGATPAAAGCTTGTGATTTCATATGAGATTGATGCTTGAGAGATGCAAG 638
OY 124 ----- 124
Db 639 GGTTTCTCCGAGAACATCAGATTGAAGCTTATCTATGACAGATGATCTTAATCT 698
OY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
    |||||||
Db 699 CGGATCCACAAAAAGTGAANTAAATGTCACTACTGCGGTTTCAGAAATGCTTGCA 758
OY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158
    |||||||
Db 759 GTGGGATGTCATATATGCCATCAGGTTTGGGCGGATCGCACAGGCCGAGAGAGAG 818
OY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluAlaAspLeuArg 178
    |||||||
Db 819 CTGTGTGGCGAGATCCAGTATATCGACAGCTCAATCCAGAGCCCTGACCTCCCT 878
OY 179 ---AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAla 197
    |||||||
Db 879 CAGGCCCTGGCAAAACATTGTATGCTCATATCAATTAATCTCTTCCCGGTGACCAAGCA 938
OY 198 LysAlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAsp 217
    |||||||
Db 939 AAGGCGAGGCGCATCTTGCAGAGAAAGACACAGACAAATCAACATTCCTTATCATATGAC 998
OY 218 MetAsnSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGln 237
    |||||||
Db 999 ATGAATTCCTTATATGAGAGAGATAAATCAATCAATCAACATTCACCCCGCTGCAG 1058
OY 238 GluGlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGlu 257
    |||||||
Db 1059 GAGCAACAGCAAGAGTGGCCATCCCATCTTTCAGAGGCTGCCAGTTTGCTCCGCGAG 1118
OY 258 AlaValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeu 277
    |||||||

```

```

Db 1119 CCTGTGACAGAGATCACAGAGTATGCCAAAAGCATTCCTGTTTGTAAATCTGACTTG 1178
OY 278 AsnAspGlnValThrLeuLeuLysTyrGluValHisGluIleIleTyrThrMetLeuAla 297
    |||||||
Db 1179 AAGGACCAAGTACTCTCTCAAAATATGGAGTCACAGAGATCATTTACCAATGCTGGCC 1238
OY 298 SerLeuMetAsnLysAspGlyValLeuIleSerGlnGluGlyPheMetThrArgGlu 317
    |||||||
Db 1239 TCCTTGATGATAAATGAATGGGTTCTCATATCCGAGGGCCAAAGCTTCTATGCAAGGCGAC 1298
OY 318 PheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAla 337
    |||||||
Db 1299 TTTCATAAAGAGCCGTGGAAGCCCTTTGGTGTACTTATGAGACCCCAAGTTGAGTTGCT 1358
OY 338 ValLysPheAsnAlaLeuGluLeuAspAspSerAspLeuAlaIlePheIleAlaValIle 357
    |||||||
Db 1359 GTGAAGTTCAAGTCAGCTGGAATTAATGATGACAGGATGGCAATTTATTCCTGTCAAT 1418
OY 358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAsp 377
    |||||||
Db 1419 ATTCACAGTGGAGACCGCCGCAAGTTTGTGATGTGGAAGCCCATGGAACATTCACAGAC 1478
OY 378 AsnLeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeu 397
    |||||||
Db 1479 AACCTGCTACAGCCCTGAGCTCCAGCTGAGCTGAAGCTGAACCATCTGAGTCTCACAGCTG 1538
OY 398 PheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGln 417
    |||||||
Db 1539 TTTGGCAAGCTCTCCAGAAATGACAGACCTTCACACAAATTTGTACGGAACACGTGCAG 1598
OY 418 LeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGlu 437
    |||||||
Db 1599 CTACTGCAGATATCAAGAAGACGAGACAGACATGAGTCTTACCCGCTCTCTGACAGAG 1658
OY 438 IleTyrLysAspLeuTyr 443
    |||||||
Db 1659 ATCTACAAGACTGTATAC 1676

```

RESULT 3

```

US-09-765-111A-22
; Sequence 22, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PARAGANNA NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 2596
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (161)...(2596)
US-09-765-111A-22

```

Alignment Scores:

Pred. No.:	2,966-237	Length:	2596
Score:	2106.50	Matches:	425
Percent Similarity:	84.45%	Conservative:	4
Best Local Similarity:	83.66%	Mismatches:	14
Query Match:	60.05%	Indels:	65
DB:	10	Gaps:	3


```

QY 18 ThrAspThrLeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetPro 37
    ::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 1159 GCGTCTCTCGCCATCGCAGCGCATGGGAGAAATGACCATGGTTCACAGACAGATGCCA 1218
QY 38 PheTTPProThrAsnPhenGlyIleSerSerValAspLeuSerValMetGluAspHisSer 57
    1219 TTTCTGGCCACCACTTTGGATCGAGTCCTCCGATCTCTCCGATATGGAACACCATCC 1278
QY 58 HisSerPheAspIleLeuSerProPheThrThrValAspPheSerSerIleSerThrProHis 77
    1279 CACTCTCTGATATATCAAGCCCTCACTACTGTTGACTCTCCAGCATTTCTACTCCACAT 1338
QY 78 TyrGluAspIleProPheThrArgThrAspProValValAlaAspTyrIleTyrAspLeu 97
    1339 TACGAAGACATTCATTCATCCAGAACAGATCCAGTGGTTCAGATATACAGATGACCTG 1398
QY 98 LysLeuGlnGluTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSer 117
    1399 AACCTTCAAGAGTACCAAAAGTCAATCAAAAGGAGCCCTGCATCTCCACTTATTATCT 1458
QY 118 GluLysThrGlnLeuTyrAsn----- 124
    1459 GAGAAGACTCAGCTCTACAAATAGCCTCATGAAGACCTTCCACTCCCTCATGSCAAT 1518
QY 124 ----- 124
    1519 GAATGTCGTCTGTGGAGATAAAGCTTCTGATTCACATGAGATTTCATCTTGTGAA 1578
QY 124 ----- 124
    1579 GGATCGAAGGTTTCTTCGCGAAGAACATCAGATTGAAGCTTATCATGACAGATGTGAT 1638
QY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLys 135
    1639 CTTAACTGTGGATCCACAAAAAAGTAAATAATGATGATGCTGTCGTTTCAGAAA 1698
QY 136 CysLeuAlaValAlaGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnLaglu 155
    1699 TGCCTTGCAGTGGGATGCTCATATGCAATGCATCAGTTGGGCGGATGCCACAGGCCGAG 1758
QY 156 LysGluLysLeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAla 175
    1759 AAGGAGAAGCTGTGGCGGAGATCTCCAGTGAATGCAACCACTGAATCCAGAGCCGCT 1818
QY 176 AspLeuArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThr 195
    1819 GACCTCCGGGCCCTGGCAAAACATTTGTATGACTCATATCAATAAAGTCTTCCGCTGACC 1878
QY 196 LysAlaLysAlaArgAlaIleLeuThrGlyLysThrAspLysSerProPheValIle 215
    1879 AAAGCAAAAGCGAGGCGCATCTTGACAGAAAGACAGACAGCAAAATCCACATTCGTTATC 1938
Db 1939 TATGACATGAAATTCCTTATGATGAGAGATAAATCAATCAACACATCACCCCC 1998
QY 216 TyrAspMetAsnSerLeuMetMetGlyLysAspLysIleLysPheLysHisIleThrPro 235
    1939 TATGACATGAAATTCCTTATGATGAGAGATAAATCAATCAACACATCACCCCC 1998
QY 236 LeuGlnGlnGlnSerLysGlnValAlaIleArgIlePheGlnGlyCysGlnPheArgSer 255
    1999 CTGCAAGGACACAGCAAAAGGTGGCCATCTTTCACAGGCTGCGCAGTTTGGCTCC 2058
QY 256 ValGluAlaValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeu 275
    2059 GTGAGAGCGGTGCAGAGATCCACAGATGCGCAAAAGCATTCCTGGTTTGTAAATCTT 2118
QY 276 AspLeuAsnAspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMet 295
    2119 GACTTGAAGACCAAGTAACTCTCTCAAAATATGAGATCCAGATCATTTACAAATG 2178
Db 2119 GACTTGAAGACCAAGTAACTCTCTCAAAATATGAGATCCAGATCATTTACAAATG 2178
QY 296 LeuAlaSerLeuMetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThr 315
    2179 CTGGCTCTCTTATGATAAATGAATGAGGTTCATATCCGAGGCGCAAGGCTTCATGACA 2238

```

```

QY 316 ArgGluPheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGlu 335
    2239 AGGAGTTTCTTAAGAGCGCTGGAAAGCCTTTGGTGAATTTATGAGACCCCAAGTTTGA 2298
QY 336 PheAlaValLysPheAsnAlaLeuGluLeuAsnAspSerAspLeuAlaIlePheIleAla 355
    2299 TTTGCTGTGAAGTTCAATGCACCTGGAATATGATGACAGAGCACTTGCAATATTTATCT 2358
QY 356 ValIleIleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIle 375
    2359 GTCATTTATTTCTAGTGGAAACCGCCAGGTTTGTGAAATGTGAAGCCCATTTGAAGACAT 2418
QY 376 GlnAspAsnLeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSer 395
    2419 CAGACAACTCTCTCAAGCCCTGGAGCTCCAGCTGGAAGCTAACACCCCTGATGCTCA 2478
QY 396 GlnLeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHis 415
    2479 CAGCTGTTGGCCAACTGCTGCTCCAGAAATGACAGACCTTCAGACATTTGTCACGGAAC 2538
QY 416 ValGlnLeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeu 435
    2539 GTGCAGCTACTGCAAGTATCAAGAAAGGAGAGACAGACATGATGATTCACCCGCTCTG 2598
QY 436 GlnGluIleTyrLysAspLeuTyr 443
    2599 CAGGAGATCTACAGAGACTTTGAC 2622
Db 2599 CAGGAGATCTACAGAGACTTTGAC 2622

RESULT 5
US-09-765-111A-1
: Sequence 1, Application US/09765111A
: Patent No. US20020106796A1
: GENERAL INFORMATION:
: APPLICANT: Fletcher, Jonathan A.
: TITLE OF INVENTION: PAX8-PPARGAMMA NUCLEIC ACID MOLECULES
: FILE REFERENCE: BOB01/7196/ERP/MAT
: CURRENT APPLICATION NUMBER: US/09/765,111A
: CURRENT FILING DATE: 2001-01-18
: PRIOR APPLICATION NUMBER: US 60/177,109
: PRIOR FILING DATE: 2000-01-20
: PRIOR APPLICATION NUMBER: US 60/225,079
: PRIOR FILING DATE: 2000-08-14
: NUMBER OF SEQ ID NOS: 47
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 2334
: TYPE: DNA
: ORGANISM: Homo Sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(2334)
: US-09-765-111A-1

Alignment Scores:
Pred. No.: 1.91e-236 Length: 2334
Percent: 2099.00 Matches: 420
Best Similarity: 84.33% Conservative: 5
Best Local Similarity: 83.33% Mismatches: 13
Query Match: 59.83% Indels: 66
DB: 10 Gaps: 2

US-09-931-007a-1 (1-688) x US-09-765-111a-1 (1-2334)
QY 6 GlyAspSerProIleAspProGluSerAspSerPheThrAspThrLeuSerAlaAsn--- 24
    820 GGGAGAGGCAACCCCTGACCCCTTCCAAACGCGACGCGGCGCAACCTCTCAGCTACACG 879
Db 820 GGGAGAGGCAACCCCTGACCCCTTCCAAACGCGACGCGGCGCAACCTCTCAGCTACACG 879
QY 25 -----IleSerGlnGluMetThrMetValAspThrGluMetProPheThrProThr 41
    880 ACCTACCCCGTGTGGCAGAAATGACATGTTGACACAGAGATGCCATTCCTGGCCGACC 939

```

OY	42	AspPhegylIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPheasp	61
Db	940	AACTTTGGGATCACCTCCGGGAGCTCTCCGTATATGAAGACACCCACCTCCATTGAT	999
OY	62	IleIysProPheThrThrValAspPheSerSerIleSerThrProHisTyrIleuAspIle	81
Db	1000	ATCAAGCCCTTCACTACTGTTGACTTCTCCAGCATTTTCTACTCCACATTACGAACACATT	1055
OY	82	ProPheThrArGthrAspProValValAlaAspTyrIleSerTyrAspLeuIysLeuGlnIu	101
Db	1060	CCATTCAACAAGACAGATCCAGTGGTTCAGATTACAAAGATATGACTGAACCTCAAGAG	1119
OY	102	TyrGlnSerAlaIleIysValGluProAlaSerProProTyrTyrSerGluIysThrGln	121
Db	1120	TACCAAACTGCAAATCAAAAGTGGACCCGTGACTTCCACCTTATATTCTGAGAAGACTGAG	1179
OY	122	LeuTyrAsn-----	124
Db	1180	CTCTTCAATTAAGCCTCATGGAAGCCTTCCAACTCCCTATGGCAATTGAATGTGTCTC	1238
OY	124	-----	124
Db	1240	TGTGGAGATAAAGCTCTGTGATTCTCACTATGAGATTTCATGCTTGGAAGATGCAAGGT	1299
OY	124	-----	124
Db	1300	TTCTTCGGAGACAAATCAGATTGAACCTTATCTATAGACAGATGTGATCTTAACGTCCG	1359
OY	125	-----ArganIysCysGlnTyrCysArgPheGlnIleCysCysLeuAlaI	139
Db	1360	ATCCACAAAAAAGTACAAATTAATGTCACTACTGTCGGTTTTCAGAAATGCCCTTCAGTG	1419
OY	140	GlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluIysGluIysLeu	159
Db	1420	GGGAGTCTCATTAATGCCATCAGGTTTGGCGGATGCCACAGGCGAGAGAGAAAGACTG	1479
OY	160	LeuAlaGluIleSerSerAspIleAspGlnIleAsnProGluSerAlaAspLeuArgAla	179
Db	1480	TTGGGGGAGATCTCACTAGTATTCGACACCTGAATCCAGAGTCGGTCACTCCGGGCC	1539
OY	180	LeuAlaIysHisIleuTyrAspSerTyrIleIysSerPheProIeuThrIysAlaIysAla	199
Db	1540	CTGGCAAAACCTTTGTATGTACTCATCATTAAGTCTTCCCGTGCACCAAGCAAAAGCG	1599
OY	200	ArgAlaIleIeuThrGlyIysThrThrAspIysSerProPheValIleTyrIysPheIasn	219
Db	1600	AGGGGATCTTGACAGAGAAAGACACAGCAACATCAACATCGGTTATGTATCATCATGAT	1659
OY	220	SerLeuMetMetGlyIuAspIysIleIysPheIysHisIleThrProLeuGlnGln	239
Db	1660	TCTTTAAAGATGGAGAAAGATAAATCTCAAGTTCAACATCAACATCACCCCTCGACGAGCAG	1719
OY	240	SerIysGluValAlaIleArgIlePheGlnIuCysGlnPheArgSerValGluAlaI	259
Db	1720	AGCAAGAGGTGGCCATCCCATCTTTCAGGGCTGCACATTCCTCCGTGGAGGCTGTG	1779
OY	260	GlnGluIleThrGluTyrAlaIysSerIleProGlyPheValAsnLeuAspLeuIysnAsp	279
Db	1780	CAGGAGATCACAGAGTATGCCCAAAAGCATTCCTCGGTTTGTAAATCTTGACTTGACACAC	1839
OY	280	GlnValThrLeuLeuIysTyrIlyValHisGluIleIleTyrThrMetLeuAlaSerIeu	299
Db	1840	CAACTTAATCTCCCAATATATGAGATCCACAGATCATTTACACATCTGTGCTCTTCYTG	1899
OY	300	MetIasnIysAspGlyValIleuIleSerGluGlyGlnGlyPheMetThrArgIuPheIeu	319
Db	1900	ATGATATTAAGTATGGGTTCTCATATCCGAGGGCCAGGCTTATATACAAGGAGATTCTTA	1959
OY	320	LysSerLeuArqIysProPheGlyAspPheMetGluProIysPheGluIlePheAlaIys	339
Db	1960	AAGAGCCTCGCAAAAGCTTTTGGTGCATTTATGAGACCCAAATTTGAGTTTGCTGTGAG	2019
OY	340	PheAsnIleLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIleIeu	359

Accession	Protein	Length
Db 2020	TTCAATGCACTGGAAATTAGATGACAGCGACTTGGCAATTTTAATTCGCTCACTTATTCTC	2079
Qy 360	SerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeu	379
Db 2080	AGTGAGAGACGGCCCAAGSTTTGTGTAATGTGAACCCCATTTGAAGACATTCAACAACCTG	2139
Qy 360	LeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPheAla	399
Db 2140	CTCAACAGCCCTGGAGCTCCAGCTGAAGCTGAACACCCTCGAGTCTCTCAACAGCTGTATTGCC	2199
Qy 400	LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeuLeu	419
Db 2200	AAGTGCTCCAGAAATATGACAGACCTTAGCAATTTGTACCGAACAACCTGCAGCTACTTG	2259
Qy 420	GlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIleLeu	439
Db 2260	CAGTGATCAAGAAGACGGAGACAGACATGAGTCTTCACCGCTCTCTGAGAGATATCTAC	2319
Qy 440	LysAspLeuTyr 443	
Db 2320	AAGGACTTGTATC 2331	

```

RESULT 6
US-09-765-111A-3
: Sequence 3, Application US/09765111A
: Patent No. US20020106796A1
: GENERAL INFORMATION:
: APPLICANT: Fletcher, Jonathan A.
: APPLICANT: Kroll, Todd G.
: TITLE OF INVENTION: PAX8-PPARgamma NUCLEIC ACID MOLECULES
: TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
: FILE REFERENCE: B0801/7196/ERP/MAT
: CURRENT APPLICATION NUMBER: US/09/765,111A
: PRIOR FILING DATE: 2001-01-18
: PRIOR APPLICATION NUMBER: US 60/177,109
: PRIOR FILING DATE: 2000-01-20
: PRIOR APPLICATION NUMBER: US 60/225,079
: PRIOR FILING DATE: 2000-08-14
: NUMBER OF SEQ ID NOS: 47
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 3
: LENGTH: 2523
: TYPE: DNA
: ORGANISM: Homo Sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1) ... (2523)
US-09-765-111A-3

Alignment Scores:
Pred. No.:      8,31e-236      Length:      2523
Score:          2094.00        Matches:      416
Percent Similarity: 87.03%    Conservative: 0
Best Local Similarity: 87.03%  Mismatches:  0
Query Match:     59.69%       Indels:       62
DB:              10           Gaps:         1

US-09-931-007A-1 (1-688) x US-09-765-111A-3 (1-2523)

QY      28  GIUwethrmetValAspThrGIUwethrProPhetThrProThAsnPhgIyleSerSer 47
      |||
Db      1087 GAATGACCATGGTTGACACAGAGATGCCATTCTGGCCACCAACATTGGGATCAGCTCC 116

QY      48  ValAspLeuSerValMetGIUAspHisSerHisSerPheAspIleLysPProPhetThr 67
      |||
Db      1147 GTGGATCTCTCCGTAATGAGACACACCTCCACCTCTTGATATCAACCCCTCATTACT 1206

QY      68  ValAspPheSerSerIleSerThrProHisThrGIUAspIleProPhetThrAArgThrAsp 87
      |||
Db      1207 GTTGACTTCTCCAGCAATTCTACTCCACATTAACGAAGACATTCCATTCCACAAGAACGAT 1266

QY      88  ProValValAlaAspPylrLysTrpAspLeuLysLeuGInGluTrpGlnSerAlaIleLys 107

```

```
|||||
Db 1267 CCGATGGTTCACATTACAGATGATGACCTGAACTTCACAGATACCAAAATGCAATCAAA 1326
QY 108 ValG1PProLaseRProproTyTySerGluYsThcGlnLeuTyAsn----- 124
Db 1327 GTGAGACCTGCATCTCCACCTTTATTCTGAGAGAGCTCAGCTTACATTAAGCTCAT 1386
QY 124 ----- 124
Db 1387 GAGAGAGCTTCCAACTCCCTCATGSCAATTGATGTGCTGTCTGTGAGATAAAGCTTCT 1446
QY 124 ----- 124
Db 1447 GGATTTCACTATGAGTGCATGCTGTGAAAGATGACAGAGCTTTCTCCGAGAACATC 1506
QY 125 -----Arg 125
Db 1507 AGATTGAGCTTCTATGACAGATGTGATCTTAAGTCCGATCCACAAAAAAGTAGA 1566
QY 126 AsnLysCysGlnTyrcysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAla 145
Db 1567 AATAAATGTGATGACTGTGGCTTTCAGAAATGCTTGACAGTGGGAGTGTCTAATATGCC 1626
QY 146 IleArgPheGlyArgMetProGlnAlaGluLysGluLysLeuAlaGluLysSerSer 165
Db 1627 ATCAGGTTTGGGGGATGCGACAGCGCCGAGAAAGAGAGAGAGCTGTTGGCGGAGATCTCCAGT 1686
QY 166 AspIleAspGlnLeuAsnProGlnSerAlaAspLeuArgAlaLeuAlaLysHisLysLeuTy 185
Db 1687 GATATGACACAGCTGATTCACAGATCCGCTGACCTCCGGCCCTGGCAAAACATTTGTAT 1746
QY 186 AspSerTyrlleYsSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGly 205
Db 1747 GACTCATACATAAAGTCCTCCCGCTGACCAAAAGCGGCGAGATCTTGACAGGA 1806
QY 206 LysThrThrAspLysSerProPheValIleTyAspMetCAsnSerLeuMetGlyGlu 225
Db 1807 AAGACAAACACAAATCACCATTCCGTATCTATGACATGATTCCTTAATGATGGAGAA 1866
QY 226 AspLysIleLeuPheLysHisIleThrProLeuGlnGlnSerLysGluValAlaIle 245
Db 1867 GATTAATAATCAGTTCAAAACATCACCCCTCCAGAGAGAGCAAAAGGGGCGCATC 1926
QY 246 ArgIlePheGlnGlnCysGlnPheArgSerValGluAlaValGlnIleThrGluTy 265
Db 1927 CGCATCTTTCAGGGCGTCCAGCTTCCTCGTGGAGGCTGTGAGGAGATCACAGATTA 1986
QY 266 AlaLysSerIleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLys 285
Db 1987 GCCAAAAGCATTCCTGGTTTGTAAATCTGACTGAACGACCAAGTAACCTCCCTCAAA 2046
QY 286 TyrG1ValHisGlnIleIleTyThrMetLeuAlaSerLeuMetAsnLysAspGlyVal 305
Db 2047 TATGAGATCCACAGATCATTTATACAAATGCTGGCTCTGTGATGAATGAAGTGGGGTT 2106
QY 306 LeuIleSerGlnGlnGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysPro 325
Db 2107 CTCATATCCAGAGGCGCAAGGCTTCATGACAAAGGAGTTCTTAAGAAGCGCTGCAAAAGCGCT 2166
QY 326 PheGlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeu 345
Db 2167 TTTGGTGTACTTATGAGGCCCAAGTTGAGTTGCTGTGAAGTTCAATGACACGGAATTA 2226
QY 346 AspAspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGly 365
Db 2227 GATGACAGCAGCATTTGCATATTTATTTGCTGATATTTCTGATGAGAGACCCCAAGGT 2286
QY 366 LeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeuGluAlaLeuGluLeu 385
Db 2287 TTGCTGAATGTGAAGCCCATTTGAAGACATTCAGACCAACCTGCTACAAAGCCCTGAGAGCTC 2246
QY 386 GlnLeuLysLeuAsnHisProGlnSerGlnLeuPheAlaLysLeuLeuGlnLysMet 405
|||||
```

```
Db 2347 CAGCTGAAGCTGAACCAACCCCTGAGTCCATCAAGCTTTTGCCAAAGCTGCTCCAGAAATG 2406
QY 406 ThrAspLeuArgGlnIleValThrGluHisValGlnLeuGlnValIleLysLysThr 425
Db 2407 ACAGACCTCGACAGCATTTGTACAGGAGACACGTCGACTACTGAGGTGATCAAGAAAGAG 2466
QY 426 GluThrAspMetSerLeuHisProLeuLeuGlnGluIleTyLysAspLeuTy 443
Db 2467 GAGACAGACATGATGCTTCCACCCGCTCGTGCAGAGATCTACAAAGACCTGTGAC 2520

RESULT 7
US-09-765-111A-26
; Sequence 26, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; TITLE OF INVENTION: PAX8-PPARgamma NUCLEIC ACID MOLECULES
; FILE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 1811
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (173)..(1609)
US-09-765-111A-26

Alignment Scores:
Pred. No.: 1,11e-232 Length: 1811
Score: 2065.50 Matches: 414
Percent Similarity: 85.71% Conservative: 6
Best Local Similarity: 84.49% Mismatches: 7
Query Match: 58.88% Indels: 63
DB: 10 Gaps: 2

US-09-931-007A-1 (1-688) x US-09-765-111A-26 (1-1811)
QY 23 AlaAsnIleSerGlnGlnIleThrMetValAspThrGluMetProPheThrProThrAsn 42
Db 155 GCCGCGGTGGCCGACAGAAATGACATGTTGACACAGAGATGCGATTTGGCCACCAAC 214
QY 43 PheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPheAspIle 62
Db 215 TTTGGGATGAGCTCCCTGGATCTCTCCGTAATGAAGACACATCCCATCTTTGATATC 274
QY 63 LysProPheThrThrValAspPheSerIleSerThrProHisTyrcGluAspIlePro 82
Db 275 AAGCCCTTCACATCTTGTGACTTCTCCAGCATTTCTACTCCACATTAACGAGACATTCGA 334
QY 83 PheThrArgThrAspProValValAlaAspTyrlLysTyAspLeuLysLeuGlnGluTy 102
Db 335 TTCACAAAGAACAGATCCAGTGGTTGCAGATTACAAAGTATGACCTGAAACTTCAAGAGTAC 394
QY 103 GlnSerAlaIleLysValGluProAlaSerProProTyTySerGluLysThrGlnLeu 122
Db 395 CAAGTGCATCAAAATGGAGCGCTGCATCTCCACCTTATTTATTTGAGAAAGACTCAGCTC 454
QY 123 TyrAsn----- 124
Db 455 TACAATTAACCTTCATGAAGAGCCTTCCAACTCCCTCATGGAATTAATGTGCTGTCTGT 514
QY 124 ----- 124
```

Db 515 GGAGATAAAGCTTCTGATTTCACTATGAGATTGCTGTTGTGAAGATGCAAGGTTTC 574
QY 124 ----- 124
Db 575 TTCCGGAGAACATCAGATTGAACCTTATCTAGACAGATGTGATCTTAAGTGTGGATC 634
QY 125 -----Argasnlyscysglntrycysarphheglnlyscysleualavalgly 140
Db 635 CACAAAAAAGTGAATAATATGCTACTGCTGCTTTCAGAAATGCTTGGAGTGGG 694
QY 141 MetSerHisasnAlaIleargpHeGlyArgMetProGlnAlaGlnlysglnlyscysleu 160
Db 695 ATGCTCATATGCTCATGAGTTTGGCGGATCCACAGCGCGGAGAGAGAACTGTTG 754
QY 161 AlaglnIleSerSerAspIleAspGlnLeuAsnProGlnSerAlaAspLeuArg---Ala 179
Db 755 GCGGAGATCTCCAGTATATCGACCACTGATCCAGAGTCGGTGACCTCCGCGCAGGCC 814
QY 180 LeuAlaIlyshisLeuTyraSpSerTyrIleLySerPheProLeuThrIlysaIalysa 199
Db 815 CTGGCAAAACATTTGTATGACTCATACATAAAGTCTTCCCGCTGACCAAGCAAGCGC 874
QY 200 ArgAlaIleLeuThrGlyLystrThrAspLySerProPheValIleTyraSpMetAsn 219
Db 875 AGGCGCATCTTGACAGGAAAGACAAACAGCAAAATCACCATTCGTTATCTATGACATGAAT 934
QY 220 SerLeuMetMetGlyGlnAspLyshIleLyshPheLyshIleThrProLeuGlnGln 239
Db 935 TCCCTTAATGATGGAGAGATAAATCAATCAACATCACATCCCTCCGCGCAGAGACAG 994
QY 240 SerLyshGlnValAlaIleArgIlePheGlnIlyCysGlnPheArgSerValGlnAlaVal 259
Db 995 AGCAAAAGGTGGCGCATCCCGCATTTTCAGGGCTGCCAGTTGCTCCGTGGAGGCTGTG 1054
QY 260 GlnGlnIleThrGlnTyraIalysSerIleProGlyPheValAsnLeuAspLeuAsnAsp 279
Db 1055 CAGGAGATCACAGAGTATGCCAAAGCATTCGTTTGTAAATCTTGACTTGAAAGCAC 1114
QY 280 GlnValThrLeuLeuLystrGlyValHisGlnIleIleTyrrhrMetLeuAlaSerLeu 299
Db 1115 CAAGTACCTCTCCCAATATGGAGTCCAGAGATCATTTACAAATGCTGGGCTCTTTC 1174
QY 300 MetAsnLyshAspGlyValLeuIleSerGlnIlyGlnIlyPheMetThrArgGlnPheLeu 319
Db 1175 ATGAATTAAGATGGGCTTCATATCCGAGGCCAAGGCTTCATGCAAGAGGATTTCTA 1234
QY 320 LySerLeuArgLyshProPheGlyAspPheMetGlnProLyshPheGlnPheAlaValLys 339
Db 1235 AAGAGCCTGGAAAGCCTTTTGGTGAATTTATGAGGCCCAAGTTTGAAGTTGCTGTGAG 1294
QY 340 PheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIleLeu 359
Db 1295 TTCAATGCTCATGAAATTAATGATGACAGGACTTGGCAATATTTATTTGCTGATTAATTC 1354
QY 360 SerGlyAspArgProGlyLeuLeuAsnValLysProIleGlnAspIleGlnAspAsnLeu 379
Db 1355 AGTGAAGACCGCCAGGTTTGTGGAATGTGAACCCCATTTGAAGACATTCGAAGCAACCTG 1414
QY 380 LeuGlnAlaLeuGlnLeuGlnLeuLyshLeuAsnHisProGlnSerSerGlnLeuPheAla 399
Db 1415 CTGCAAGCCCTGGAGCTCCCGCTGAAGCTGAACACCCTGAGTCTCACAGCTGTTGGCC 1474
QY 400 LyshLeuLeuGlnLyshMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeuLeu 419
Db 1475 AAGCTGCTCCAGAAATGAGAGACTCAGACAAATTTGTCAGGAACACGTCGAGCTACTG 1534
QY 420 GlnValIleLyshLystrGlnTyraSpMetSerLeuHisProLeuLeuGlnGlnIlyLeuTy 439
Db 1535 CAGGTGATCAAGAAAGAGAGACAGACATGAGTCTTACCCGCTCCGTCGAGAGATCTTAC 1594
QY 440 LysAspLeuTyraIalThrAlaIleLeuThr 449
Db 1595 AAGGACTTGTACTAGCAGAGAGTCTGTGAGC 1624

RESULT 8
US-09-880-107-2212
; Sequence 2212, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2212
; LENGTH: 1811
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 LA0904
US-09-880-107-2212

Alignment Scores:
Pred. No.: 1,11e-232 Length: 1811
Score: 2065.50 Matches: 414
Percent Similarity: 85.71% Conservative: 6
Best Local Similarity: 84.49% Mismatches: 7
Query Match: 58.88% Indels: 63
DB: 10 Gaps: 2

US-09-931-007A-1 (1-688) x US-09-880-107-2212 (1-1811)
QY 23 AlaAsnIleSerGlnGlnMetThrMetValAspThrGlnMetProPheThrProThrAsn 42
Db 155 GCGCGCGTGGCCGCGCAAGAAAGACATGTTGACACAGAGATGCTCTGCGCCACCAC 214
QY 43 PheGlyIleSerSerValAspLeuSerValMetGlnAspHisSerHisSerPheAspIle 62
Db 215 TTTCGATGATCGCTCCGAGATCTCTCGATATGGAAGACACCTCCACCTCTTGATATC 274
QY 63 LysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGlnAspIlePro 82
Db 275 AAGCCCTTCACTACTGTTGACTTCTCCAGCATTTCTACTCCACATTAACGAAGACATTCCA 334
QY 83 PheThrArgThrAspProValValAlaAspTyrLystrAspLeuLyshLeuGlnIlyTy 102
Db 335 TTCAACAAGACAGATCCAGTGGTTGCGAGATTACAAAGATACCTGGAACCTCAAGAGTAC 394
QY 103 GlnSerAlaIleLyshValGlnProAlaSerProProTyrTyrSerGlnLystrGlnLeu 122
Db 395 CAAGAGCATATCAAGATGGAGGCTGCATCCACTTATATTTCTGAGAAGACTCAGCTC 454
QY 123 TyrAsn----- 124
Db 455 TACAATAAGCCTCATGAAGAGCCTTCCAACTCCCTCATGGAATTAATGTCGTGTGTG 514
QY 124 ----- 124
Db 515 GGAGATAAAGCTTCTGATTTCACTATGAGATTGCTGTTGTGAAGATGCAAGGTTTC 574
QY 124 ----- 124
Db 575 TTCCGGAGAACATCAGATTGAACCTTATCTAGACAGATGTGATCTTAAGTGTGGATC 634
QY 125 -----Argasnlyscysglntrycysarphheglnlyscysleualavalgly 140
Db 635 CACAAAAAAGTGAATAATATGCTACTGCTGCTTTCAGAAATGCTTGGAGTGGG 694

OY	141	MetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluGlySerLeuLeu	160
OY	141	MetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluGlySerLeuLeu	160
Db	695	ATGTCATATATGCAATCAGATTGTGGCGGATCGCACAGGCGAGAGAAAGCTGTTG	754
OY	161	AlaGluIleSerSerAspIleAspGlnLeuAspProGluSerAlaAspLeuArg---Ala	179
Db	755	GGGGAGATCTCCAGTGTATTCAGACAGCTGAATCCAGAGCTCGCTGACCTCCGTAGCGC	814
OY	180	LeuAlaLysHisLeuIleYrAspSerTyrIleLysSerPheProLeuThrLysAlaLysAla	199
Db	815	CTGGCAAAACATTTGTTATGACTCATACATTAAGTCTCTCCGCTGACCAACAAAGGCG	874
OY	200	ArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMetAsn	219
Db	875	AGGGCGATCTTGACAGGAAAGACAAACAGACAATTCACATTCGTTATCTATGTACATGAAT	934
OY	220	SerLeuMetMetGlyLysAspLysIleLysPheLysHisIleThrProLeuGlnGln	239
Db	935	TCTTTATATGATGGAGAGATTAATAATCAAGTTCAACACATCAACCCCTCGACGAGACAG	994
OY	240	SerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluValAla	259
Db	995	AGCAAGAGGTGGCCATCCGCACTTTTCAGGGCTGGCAGTTTCGCTCGTGAGGCTGTG	1054
OY	260	GlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsnAsp	279
Db	1055	CAGAGATATCACAGATGTGCAAAACACATCTCGTTTGTAAATCTTGACTTGAACAC	1114
OY	280	GlnAlaThrLeuLeuLysTyrGlyValHisGlnIleIleTyrThrMetLeuLysSerLeu	299
Db	1115	CAAGTAACTCTCCCAATATTTGAGATGCCAGATATTTTAAACAAATGCTGGCTCTTG	1174
OY	300	MetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGluPheLeu	319
Db	1175	ATGATATAAGATGGGCTTTCATATCCAGAGGCCCAAGGCTTCATGCAAGGAGTTTCTA	1234
OY	320	LysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaValLys	339
Db	1235	AAGAGCTCGCAAAAGCCCTTTTGTGTGATTTATGTGAGCCCAAGTTTGAGTTTCTCTGAAG	1294
OY	340	PheAsnAlaLeuGlnLeuAsnAspSerAspLeuAlaIlePheIleAlaValIleLeuLeu	359
Db	1295	TTCAATGACATCGAATTAGATGACAGCGACTTGGCAATTTTATTTGCTGTCTATTCTTC	1354
OY	360	SerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeu	379
Db	1355	AGTGAGAGCCGCCAGGTTGTGTAATGTGAAACCCATTTGAAGACATTCMAACAACTCG	1414
OY	380	LeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPheLea	399
Db	1415	CTACAGAGCCCTGGAGCTTCCAGCTGAAGCTGAACACCCGATGATCTCCACAGCTGTTTGGC	1474
OY	400	LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeuLeu	419
Db	1475	AAGCTGCTCCAGAAATAGACAGACCTCAGACAAATTTGTACGGAACAGCTGACGCTACTG	1534
OY	420	GlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIleTyr	439
Db	1535	CAGGTGATCAAGAAAGACGGAGACAGACATGAGTCTTCCACCCGCTCGACGAGATCTTAC	1594
OY	440	LysAspLeuTyrAlaThrAlaIleLeuThr	449
Db	1595	AAGGACTTGTACTAGCAGAGAGTCTGAGC	1624
RESULT 9			
US-09-816-828-8/c			
: Sequence 8, Application US/09816828			
: Patent No. US20020150898A1			
: GENERAL INFORMATION:			
: APPLICANT: Tang, Y. Tom			
: APPLICANT: Zhou, Ping			
: APPLICANT: Goodrich, Ryle			

```

? APPLICANT: Asundi, Vinod
? APPLICANT: Ren, Feiyan
? APPLICANT: Xue, Aidong J.
? APPLICANT: Ma, Yunqing
? APPLICANT: Wang, Zhiwei
? APPLICANT: Zhao, Qing A.
? APPLICANT: Zhang, Jie
? APPLICANT: Wang, Jian-Rui
? APPLICANT: Drmanac, Radje T.
? TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
? TITLE OF INVENTION: Polypeptides
? FILE REFERENCE: 791C1P2E
? CURRENT APPLICATION NUMBER: US/09/816, 828
? CURRENT FILING DATE: 2001-03-22
? PRIOR APPLICATION NUMBER: 09/770,160
? PRIOR FILING DATE: 2001-01-26
? PRIOR APPLICATION NUMBER: 09/552, 929
? PRIOR FILING DATE: 2000-04-18
? NUMBER OF SEQ ID NOS: 18
? SOFTWARE: pl_FL_genes Version 2.0
? SEQ ID NO 8
? LENGTH: 2260
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1612)..(2142)
? US-09-816-828-8

```

Alignment Scores:

Pred. No.:	1.56e-232	Length:	2260
Score:	85.5% 2065.50	Matches:	414
Percent Similarity:	85.71%	Conservative:	6
Best Local Similarity:	84.49%	Mismatches:	7
Query Match:	58.88%	Indels:	63
DB:	10	Gaps:	2

US-09-931-007A-1 (1-688) x US-09-816-828-8 (1-2260)

QY 23 ALAsnILesErGInglumetThrmValAspThrgIumetProphetrProThrAsn 42

Db 1657 GCCGCCGTGGCCGAGAAATGACCATGGTTGACACAGAGATCGCATTTCTGGCCCAAC 1598

43 pHeGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPheAspIle 62

D6 1597 TTTGGGATCAGCTCCGTCGATCTCTCCGTAATGGAAGACCACCTCCCACTCCTTGATATC 15

63 LysProPheIhrIhrValAspPheSerSerLeSerThrProIstIyrGluAspLePro 82

DB 153 / AAGCCCTTCACCTACGCTGACTTCTCCAGCACTTCTACTCCACATTAAGAGACATTCCTCA 14

[illegible][illegible][illegible]

— — — — —

1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1221 1222 1223 1224 1225 1226 1227 1228 1229 1230 1231 1232 1233 1234 1235 1236 1237 1238 1239 1240 1241 1242 1243 1244 1245 1246 1247 1248 1249 1250 1251 1252 1253 1254 1255 1256 1257 1258 1259 1260 1261 1262 1263 1264 1265 1266 1267 1268 1269 1270 1271 1272 1273 1274 1275 1276 1277 1278 1279 1280 1281 1282 1283 1284 1285 1286 1287 1288 1289 1290 1291 1292 1293 1294 1295 1296 1297 1298 1299 1300 1301 1302 1303 1304 1305 1306 1307 1308 1309 1310 1311 1312 1313 1314 1315 1316 1317 1318 1319 1320 1321 1322 1323 1324 1325 1326 1327 1328 1329 1330 1331 1332 1333 1334 1335 1336 1337 1338 1339 1340 1341 1342 1343 1344 1345 1346 1347 1348 1349 1350 1351 1352 1353 1354 1355 1356 1357 1358 1359 1360 1361 1362 1363 1364 1365 1366 1367 1368 1369 1370 1371 1372 1373 1374 1375 1376 1377 1378 1379 1380 1381 1382 1383 1384 1385 1386 1387 1388 1389 1390 1391 1392 1393 1394 1395 1396 1397 1398 1399 1400 1401 1402 1403 1404 1405 1406 1407 1408 1409 1410 1411 1412 1413 1414 1415 1416 1417 1418 1419 1420 1421 1422 1423 1424 1425 1426 1427 1428 1429 1430 1431 1432 1433 1434 1435 1436 1437 1438 1439 1440 1441 1442 1443 1444 1445 1446 1447 1448 1449 1450 1451 1452 1453 1454 1455 1456 1457 1458 1459 1460 1461 1462 1463 1464 1465 1466 1467 1468 1469 1470 1471 1472 1473 1474 1475 1476 1477 1478 1479 1480 1481 1482 1483 1484 1485 1486 1487 1488 1489 1490 1491 1492 1493 1494 1495 1496 1497 1498 1499 1500 1501 1502 1503 1504 1505 1506 1507 1508 1509 1510 1511 1512 1513 1514 1515 1516 1517 1518 1519 1520 1521 1522 1523 1524 1525 1526 1527 1528 1529 1530 1531 1532 1533 1534 1535 1536 1537 1538 1539 1540 1541 1542 1543 1544 1545 1546 1547 1548 1549 1550 1551 1552 1553 1554 1555 1556 1557 1558 1559 1560 1561 1562 1563 1564 1565 1566 1567 1568 1569 1570 1571 1572 1573 1574 1575 1576 1577 1578 1579 1580 1581 1582 1583 1584 1585 1586 1587 1588 1589 1590 1591 1592 1593 1594 1595 1596 1597 1598 1599 1600 1601 1602 1603 1604 1605 1606 1607 1608 1609 1610 1611 1612 1613 1614 1615 1616 1617 1618 1619 1620 1621 1622 1623 1624 1625 1626 1627 1628 1629 1630 1631 1632 1633 1634 1635 1636 1637 1638 1639 1640 1641 1642 1643 1644 1645 1646 1647 1648 1649 1650 1651 1652 1653 1654 1655 1656 1657 1658 1659 1660 1661 1662 1663 1664 1665 1666 1667 1668 1669 1670 1671 1672 1673 1674 1675 1676 1677 1678 1679 1680 1681 1682 1683 1684 1685 1686 1687 1688 1689 1690 1691 1692 1693 1694 1695 1696 1697 1698 1699 1700 1701 1702 1703 1704 1705 1706 1707 1708 1709 1710 1711 1712 1713 1714 1715 1716 1717 1718 1719 1720 1721 1722 1723 1724 1725 1726 1727 1728 1729 1730 1731 1732 1733 1734 1735 1736 1737 1738 1739 1740 1741 1742 1743 1744 1745 1746 1747 1748 1749 1750 1751 1752 1753 1754 1755 1756 1757 1758 1759 1760 1761 1762 1763 1764 1765 1766 1767 1768 1769 1770 1771 1772 1773 1774 1775 1776 1777 1778 1779 1780 1781 1782 1783 1784 1785 1786 1787 1788 1789 1790 1791 1792 1793 1794 1795 1796 1797 1798 1799 1800 1801 1802 1803 1804 1805 1806 1807 1808 1809 1810 1811 1812 1813 1814 1815 1816 1817 1818 1819 1820 1821 1822 1823 1824 1825 1826 1827 1828 1829 1830 1831 1832 1833 1834 1835 1836 1837 1838 1839 1840 1841 1842 1843 1844 1845 1846 1847 1848 1849 1850 1851 1852 1853 1854 1855 1856 1857 1858 1859 1860 1861 1862 1863 1864 1865 1866 1867 1868 1869 1870 1871 1872 1873 1874 1875 1876 1877 1878 1879 1880 1881 1882 1883 1884 1885 1886 1887 1888 1889 1890 1891 1892 1893 1894 1895 1896 1897 1898 1899 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910 1911 1912 1913 1914 1915 1916 1917 1918 1919 1920 1921 1922 1923 1924 1925 1926 1927 1928 1929 1930 1931 1932 1933 1934 1935 1936 1937 1938 1939 1940 1941 1942 1943 1944 1945 1946 1947 1948 1949 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025

α_7	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
------------	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

Db 1237 TTCCGGAACAATCAGATTTGAAGCTTATTCATATGACAGATGTCATCTTAACTGTCGGATC 11

0V 125 -----ArqAsnLysCysGlnTyrCysAroPheGlnLysCysLeuAlaValGly 14

```
Db 1177 CACAAAAAGTAGAATAAATGTCTAGTCTGCGTTTCAGAAATGCTTCGACGTGGG 1118
Qy 141 MetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluArgLysLeuLeu 160
Db 1117 ATGTCTCATTAATGCCATCAGGTTTGGCGGATCGCACAGGCCAGCAAGAGGAGGAGCTGTG 1058
Qy 161 AlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg---Ala 179
Db 1057 GCGGAGATCTCCAGTGATATCGACAGCTGAATCCAGAGTCCGCTCAGCTCCGTCAGGCC 998
Qy 180 LeuAlaLeuHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLysAla 199
Db 997 CTGGCAAAACATTTGTATGACTCATCATATAAGTCCCTCCGCTGACCAAGCAAGGCGG 938
Qy 200 ArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMetAsn 219
Db 937 AGGGCATCTTGACAGCAAGAACACACAGACAAATCCATTGCTTATCTATGACATGAAT 878
Qy 220 SerLeuMetMetLysGluAspLysIleLysPheLysHisIleThrProLeuGlnGluGln 239
Db 877 TCCTTAATGATGGAGAAATATAATCAAGTTCAAACACATCACCCCTCGACGAGACAG 818
Qy 240 SerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAlaVal 259
Db 817 ACCAAGAGGTGGCCATCCGCACTTTTCAGGGCTGCCAGTTGCTCCGCTGGAGGCTGTG 758
Qy 260 GlnGluIleThrGluTyrAlaLysSerIleProGluPheValAsnLeuAspLeuAsnAsp 279
Db 757 CAGGAAATCACAGAGATATGCCAAAGCATTCCTCGGTTTGTAAATCTTGACCTTGAACGAC 698
Qy 280 GlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeu 299
Db 697 CAAGTAACCTCTCTCAAAATATGAGTCCACAGATCATTTACACATGCTGGGCTCTCTTG 638
Qy 300 MetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgLysPheLeu 319
Db 637 ATGAATAAAGATGGGTCTCATATCCGAGGGCCAGGCTTCATGTCAAGGGAGTTTCTA 578
Qy 320 LysSerLeuArgLysProPheGlnAspPheMetGluProLysPheGluPheAlaValLys 339
Db 577 AAGAGCTTCGGAAGCCTTTTGGTGTATATGAGGCCCAAGTTTGAAGTTTGGTGTGAAG 518
Qy 340 PheAsnAlaLeuGlnLeuAspSerSerAspLeuAlaIlePheIleAlaValIleIleLeu 359
Db 517 TTCAATGACACTGGAAATTAATGATGACAGCTGGCAATATTATATCTGTCATATTATTC 458
Qy 360 SerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeu 379
Db 457 AGTGGAGACCGCCAGGTTTGTGATGTGAAGCCCATTTGAAGACATTTCAAGCAACACTG 398
Qy 380 LeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPheAla 399
Db 397 CTACAAGCCCTGGAGCTCAGCTGGAAGCTGAACCAACCCCTGAGAGCTTCACAGCTTTGCC 338
Qy 400 LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeuLeu 419
Db 337 AAGCTCTCTCCAAATAATGACAGACCTCAGACAGATTTGTACAGGAACACGTCGACGTA 278
Qy 420 GlnValIleLysLysThrGlnThrLysPheMetSerLeuHisProLeuLeuGlnGluIleTyr 439
Db 277 CAGGTGATCAAGAGAGCGAGACAGACATGAGTCTTCCACCCGCTCTCGAGGAGATCTAC 218
Qy 440 LysAspLeuTyrAlaTyrAlaIleLeuThr 449
Db 217 AAGGACTTGTACTAGCAGAGAGACTCTTGAGC 188

RESULT 10
US-10-125-237-5/c
; Sequence 5, Application US/10125237
; Publication No. US20030022329A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
```

```
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aodong J.
; APPLICANT: Xue, Jie
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radote T.
; TITLE OF INVENTION: No. US20030022329A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/125,237
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 09/668,317
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/552,929
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: pl_FL_genes Version 2.0
; SEQ ID NO 5
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1612)..(2184)
US-10-125-237-5

Alignment Scores:
Pred. No.: 1,61e-232 Length: 2307
Score: 2065.50 Matches: 414
Percent Similarity: 85.71% Conservative: 6
Best Local Similarity: 84.49% Mismatches: 7
Query Match: 58.88% Indels: 63
DB: Gaps: 2

US-09-931-007A-1 (1-688) x US-10-125-237-5 (1-2307)
Qy 23 AlaAsnIleSerGlnLysMetThrMetValAspThrGluMetProPheThrProThrAsn 42
Db 1657 GCCGCGGTGGCCGACGAATAAGACATGTTGACACAGAGATGCAATTCGGGCCACCAAC 1598
Qy 43 PheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPheAspIle 62
Db 1597 TTTGGATAGCTCCGTGATCTCTCCGTAATGGAAGACACTCCACCTCTTGTATATC 1538
Qy 63 LysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAspIlePro 82
Db 1537 AAGCCTTCACTAAGTGTGACTTCTCCACGATTTCTACCTCCACATTTACGAAAGACATTCCA 1478
Qy 83 PheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGlnGluTyr 102
Db 1477 TTTCACAGACAGATCCATGATGTTGGCATTTACAGATGACCTTCAAACTTCAAGAGTAC 1418
Qy 103 GlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThrGlnLeu 122
Db 1417 CAAAGTCAATCAAGTGAAGTGAAGCTGCATCTCACCCTTATATTTCGAGAGACTCAGCTC 1358
Qy 123 TyrAsn----- 124
Db 1357 TACAATTAAGCTTCATGAGAGCCCTTCAACCTCTCATGCAATTGAATGTCGTGTCTGT 1298
Qy 124 ----- 124
Db 1297 GGAGATAAAGCTTCGATTTTCATATGAGATTCATGCTTGTGAAGATGCAAGGCTTTC 1238
Qy 124 ----- 124
Db 1237 TTCCGAGACACATCAGATTGAAGCTTATCTATGACAGATGTCATTTAAGCTGCGGATC 1178
Qy 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGly 140
|||||
```


Db 1177 CACAAAAAAGTAGAATAAATGTCAGTACTGCTGCTTTCAGAAATGCTTGCAGTGGG 1118
Qy 141 MetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLysLeuLeu 160
Db 1117 ATGTCTCATTAATGCCATCAGGTTTGCGGAGATCGACAGCGCGAGAGAGAGAGTGTGTTG 1058
Qy 161 AlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg--Ala 179
Db 1057 GCGGAGATCTCCAGATGATGACAGCAGTGAATCCAGATCCGCTGACCTCCGTCAGGCC 998
Qy 180 LeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLysAla 199
Db 997 CTGGCAAAACATTGTATGACCATCATCAATAAAGTCTTCCGCTGACCAAAAGGCG 938
Qy 200 ArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMetAsn 219
Db 937 AGGGGATCTTGACAGGAAGAAGAACACACAAATACCATTCGTTATGTAGCACTGAT 878
Qy 220 SerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGluGln 239
Db 877 TCCTTAATGATGGAGAGATTAATCAAGTTCAAAACATCACCCCTCGAGAGAGCAG 818
Qy 240 SerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAlaVal 259
Db 817 AGCAAAAGAGGTGGCCATCCGATCTTTCAGGGCTCCAGATTCCGCTCGTGAGGCTGTG 758
Qy 260 GlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAspLeuAspLeuAsnAsp 279
Db 757 CAGGAGATCACAGAGTATGCCAAAGCATTTCTGTTTGTAACTTGACTTGACAGCAG 698
Qy 280 GlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeu 299
Db 697 CAAGTAACCTCTCCAAATATGAGATCCACAGATCATTTACAAATGCTGGCCCTCTTG 638
Qy 300 MetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGluPheLeu 319
Db 637 ATGAATAAAGATGGGGTCTCATATCCAGGGCCCAAGGCTTCATACAAAGGAGTTTCTA 578
Qy 320 LysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaValLys 339
Db 577 AAGAGCCCTGGCAAGCCCTTTGGTGACTTTATGAGCCCAAGTTTGAAGTTGCTGTGAG 518
Qy 340 PheAsnAlaLeuGlnLeuAspSerAspLeuAlaIlePheIleAlaValIleIleLeu 359
Db 517 TTCATATGACATGGAATGATGACAGCGACTTGGCAATATTATGCTCATTTTCTC 458
Qy 360 SerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeu 379
Db 457 AGTGAGACCGCCCGAGTTTGTGTAATGTGAAGCCCATTTGAAGACATTCAGACAACTG 398
Qy 380 LeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPheAla 399
Db 397 CTACAAAGCCCTGGAGCTCCAGCTGAGCTGAACCCCTGAGTCTCCACAGCTGTTTCC 338
Qy 400 LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValIleThrGlnHisValGlnLeuLeu 419
Db 337 AAGCTGCTCCAGAAATATACAGACTCAGACAGATGTGCAGGAGACAGTGCAGCTACTG 278
Qy 420 GlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIleTyr 439
Db 277 CAGGATCAAGAAACAGAGACAGACATGAGTCTTCAACCGCTCTCCAGAGATGATCTAC 218
Qy 440 LysAspLeuTyrAlaIleThrAlaIleLeuThr 449
Db 217 AAGGACTTGTACTAGCAGAGAGTCTCTGAGC 188

RESULT 11
US-09-816-828-9/c

; Sequence 9, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aldong J.
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Zhilwei
; APPLICANT: Zhao, Qing A.
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radolje T.
; TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/816,828
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/552,929
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: pt_fl_genes Version 2.0
; SEQ ID NO 9
; LENGTH: 2329
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1612)..(2211)
; US-09-816-828-9

Alignment Scores:
Pred. No.: 1,63e-232 Length: 2329
Score: 2065.50 Matches: 414
Percent Similarity: 85.71% Conservative: 6
Best Local Similarity: 84.49% Mismatches: 7
Query Match: 58.88% Indels: 63
DB: Gaps: 2

US-09-931-007a-1 (1-688) x US-09-816-828-9 (1-2329)

Qy 23 AlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrProThrAsn 42
Db 1657 GCCGCCGTGGCGCGAGAAATGACCATGTTGACACAGAGATCGATTCGTGCCCCAAC 1598
Qy 43 PheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPheAspIle 62
Db 1597 TTTGGGATCAGCTCCGTGATCTCTCCGTAATGGAAGACCACTCCCATCTTGATATTC 1538
Qy 63 LysProPheThrThrValAspPheSerIleSerThrProHisTyrGluAspIlePro 82
Db 1537 AAGCCCTTCACTAGCTGTGACTTCTCCAGCATTTCTACTCCACATTAAGAAACATTCCA 1478
Qy 83 PheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGlnGluTyr 102
Db 1477 TTCAACAAGAACAGATCCAGCTGTTGCAGATTACAGTATGACCTGAACCTCAACAGTAC 1418
Qy 103 GlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThrGlnLeu 122
Db 1417 CAAGTGCATTAAGATGAGAGCTGATCTCCACCTTATTTATTCGAGAAAGACTAGCTC 1358
Qy 123 TyrAsn----- 124
Db 1357 TACATATAAGCCTCATGAAGAGCTTCAACATCCCTCATGGCAATTGAATGTGCTGTCTG 1298
Qy 124 ----- 124
Db 1297 GGAGATAAGCTTCTGATTTACATATGAGATGATGCTTGTGAAGAGATGCAAGGGTTTC 1238
Qy 124 ----- 124
Db 1237 TTCCGGAGAACATATCAGATTGAGCTTATCTATGACAGATGATCTTAACATGCGGATC 1178
Qy 125 -----AtGAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGly 140

```
Db 1177 CACAAAAAGTGAATATCTCAGTACTCGGTTTCAGAAAGCCCTGCAGTGGGG 1118
Oy 141 MetSerHisasnAlaIleArgPheGlyArgMetProGlnAlaIleArgGlyLysLeuLeu 160
Db 1117 ATCTTCATATATCCATCCAGTTTGCGGATCGCACAGGCCAGGAAGGAGAGCTGTG 1058
Oy 161 AlaGluIleSerSerAspIleAspGlnLeuAsnProGlnSerAlaAspLeuArg---Ala 179
Db 1057 GCGGAGATCTCCAGTGATATCCAGCCAGTGAATCCAGAGTCCGCTGACCTCGTCAAGGCC 998
Oy 180 LeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLysAla 199
Db 997 CTGGCAAAACATTGTATGACTCATCATATAAAGTCCCTCCGCTGACCAAGCAAGGCGG 938
Oy 200 ArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMetAsn 219
Db 937 AGGGCATCTTGACAGGAAGACACAGACAAATCCACATTCGTTATCTATGACATGAAT 878
Oy 220 SerLeuMetMetGlyLysAspLysIleLysPheLysHisIleThrProLeuGlnGln 239
Db 877 TCCTTATGATGGGAGAGATATAATCAAGTTCAAAACATCAACCCCTGCAGAGAGAG 818
Oy 240 SerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAlaVal 259
Db 817 AGCAAGAGGTGGCCATCCGCACTTTCAGGGCTCCAGTTTCGCTCCGCGAGGCTGTG 758
Oy 260 GlnGluIleThrGlnTyrLysSerIleProGlnPheValAsnLeuAspLeuAsnAsp 279
Db 757 CAGGAATACACAGATATGCCAAAGACATTCCTGGTTTGTAAATCTGTGCTTGAAGCAG 698
Oy 280 GlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeu 299
Db 697 CAAGTAACCTCTCTCAAAATATGAGATCCAGAGATCATTTACACATATGCTGGCTCTTG 638
Oy 300 MetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGluPheLeu 319
Db 637 ATGAAATAAAGATGGGTCTCATATCCGAGGGCAAGGCTTCATGTCAAGGGAGTTCTTA 578
Oy 320 LysSerLeuArgLysProPheGlnAspPheMetGluProLysPheGlnPheAlaValLys 339
Db 577 AAGAGCTCGCAAGACCTTTTGGTGCATTTATGAGCCCAAGTTTATGTTGCTGTGAAG 518
Oy 340 PheAsnAlaLeuGlnLeuAspSerAspLeuAlaIlePheIleAlaValIleIleLeu 359
Db 517 TTCAATGCACTGGAATATGATGACAGCACTTGGCAATATTTATTCCTGCTCATTTATTC 458
Oy 360 SerGlyAspArgProGlnLeuAsnValLysProIleGlnAspIleGlnAspAsnLeu 379
Db 457 AGTGGAGACCGCCAGGTTTGGTGAATGTGAAGCCCATTTGAAGACATTCACAGCAACCTG 398
Oy 380 LeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGlnUserSerGlnLeuPheAla 399
Db 397 CTACAAGCCCTGGAGAGTCCAGCTGAAGCTGAACACCCCTGATGCTCCACAGCTGTTTGGC 338
Oy 400 LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeuLeu 419
Db 337 AAGCTGCTCCAGAAATGACAGACCTCAGACAGATGTGTACAGGAACAGCTGCAGTACTG 278
Oy 420 GlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuGlnGlnIleTyr 439
Db 277 CAGGTGATATCAAGAGAGGAGACAGATGAGTCTTCACCCCTCCTGCAGAGAGATCTAC 218
Oy 440 LysAspLeuTyrAlaTyrAlaIleLeuThr 449
Db 217 AAGGACTTGTACTAGCAGAGAGTCCCTGAGC 188
```

```
RESULT 12
US-10-142-373-1
; Sequence 1, Application US/70142373
; Patent No. US20020137665A1
; GENERAL INFORMATION:
; APPLICANT: EVANS, Ronald
```

```
; APPLICANT: FORMAN, Barry
; TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR ACTIVATED
; TITLE OF INVENTION: RECEPTOR-GAMMA.
; TITLE OF INVENTION: AND METHODS FOR THE USE THEREOF
; FILE REFERENCE: SALK1480-2
; CURRENT APPLICATION NUMBER: US/10/142,373
; PRIOR FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: US/09/788,070
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/955,302
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2005
; TYPE: DNA
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (352)..(1776)
US-10-142-373-1

Alignment Scores:
Pred. No.: 1,48e-232 Length: 2005
Score: 2065.00 Matches: 408
Percent Similarity: 86.40% Conservative: 5
Best Local Similarity: 85.36% Mismatches: 3
Query Match: 58.87% Indels: 62
DB: Gaps: 1

US-09-931-007a-1 (1-688) x US-10-142-373-1 (1-2005)
Oy 28 GluMetThrMetValAspThrGluMetProPheThrProThrAsnPheGlyLysSer 47
Db 343 GAATATACCATGATGTTGACACAGAGATGCCATTCCTGCGCCACCAATTCGGAATACGCT 402
Oy 48 ValAspLeuSerValMetGlnAspHisSerHisSerPheAspIleLysProPheThrThr 67
Db 403 GTGACCTCTCCGTGATGAGAACACCTGCAATCTCTTGTGACATCAAGCCCTTACACA 462
Oy 68 ValAspPheSerSerIleSerThrProHisTyrGlnAspIleProPheThrArgThrAsp 87
Db 463 GTTGATTTCTCCAGCATTTCTGCTCCACACTATGAAAGACATTCATTCACAGAGCTGAC 522
Oy 88 ProValAlaIleAspTyrLysTyrAspLeuLysLeuGlnGlyTyrGlnSerAlaIleLys 107
Db 523 CCAATGCTGTCGATTACAAATATGACCTGAAGCTCCAAAGATACCAAGTCCGATCAA 582
Oy 108 ValGluProAlaSerProProTyrTyrSerGlnLysThrGlnLeuTyrAsn----- 124
Db 583 GTAGAACCCTGCATCTCCACCTTATTATTCTGAAGACCCACCTCTACAAACAGGCTCAT 642
Oy 124 ----- 124
Db 643 GAAGAACCTTCTAATCTCCCTATGGCCATTTAGTGCAGTCTGTGGGATTAAGCATCA 702
Oy 124 ----- 124
Db 703 GGCTTCCACTATGAGATTCATGCTGTGAGAGATCCAAAGGTTTTTCCGAAGAACATC 762
Oy 125 -----Arg 125
Db 763 CGATTGAACCTATTATTATGATAGTGTGATCTTAATCGCGGATCCAAAAAAGTGA 822
Oy 126 AsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAla 145
Db 823 AATTAATGTCAGTACTGTGCGTTTCAGAGTGCCTTGTGCGGATGTCTCACATGCC 882
Oy 146 IleArgPheGlyArgMetProGlnAlaGlnLysGlnLysLeuLeuAlaGlnIleSerSer 165
Db 883 ATCAGGTTTGGGCGATGTCACAGGCCCAAGAGAGAAAGACTGTTTGGCGAGATCTCCAGT 942
Oy 166 AspIleAspGlnLeuAsnProGlnSerAlaAspLeuArgAlaLeuAlaLysHisLeuTyr 185
```

```

|||||
Db 943 GATATGACCAAGCTGACCCAGAGTCTGATCTGCAGACCCCTTGCAAGACTTTGTAT 1002
|||
Qy 186 ASPserTYrILeYSerPheProLeuThrLYsAlaLYsAlaArgAlaILEuThrGLY 205
|||
Db 1003 GACTCATACATTAATCTCTCCGCTGACCAAGCCAAAGCCAGGCGCATCTTGACAGGA 1062
|||
Qy 206 LysThrThrAspLYSerPheProPheValILEtyrAspMetAsnSerLeuMetGLYGLU 225
|||
Db 1063 AAGACACGAGAACAAATCAACATTTGTCTACGACATGAATGCTTATATGATGGAGAA 1122
|||
Qy 226 AspLYsILELYsPheLYsHisILEthrProLeuGLNGLNGLNserLYsGLYValaILE 245
|||
Db 1123 GATTAATAATCAAGTTCAACATATCACCCCTCGACGAGCAGACAAAGAGGTGGCCATC 1182
|||
Qy 246 ArgILEPheGLNGLYCYsGLNpHeArgSerValGLUValaGLNILEthrGLYTYR 265
|||
Db 1183 CGAATTTTCAAGGGTGCAGTTTGATCCGTGAGAACCGGTGCAAGAGATCACAGATAT 1242
|||
Qy 266 AlaLYSerILEProGLYpHeValAsnLeuAspLeuAsnAspGLNValThrLeuLeuLYs 285
|||
Db 1243 GCCAAATAATATCCCTGGTTTCACTTAACCTGATTTGAATGACCAAGTACCTGTGCTCAAG 1302
|||
Qy 286 TYrGLYValaHISGLNILEtyrThrMetLeuAlaSerLeuMetAsnLYsAspGLYVal 305
|||
Db 1303 TATGGTGTCCATGAGATCATCTACACAGATGCGCCCTCCCTGATGAATAAGATGAGATC 1362
|||
Qy 306 LeuILESerGLNGLYGLNGLYpHeMetThrArgGLNpHeLeuLYsSerLeuArgLYsPro 325
|||
Db 1363 CTCATCTCAGAGGCGCCAAAGATTCATGACCCAGGAGTTCTTAAAGCCCTGGGAGAGCC 1422
|||
Qy 326 PheGLYAspPheMetGLNProLYsPheGLNpHeAlaValLYsPheAsnAlaLeuGLNLeu 345
|||
Db 1423 TTGGTGACTTATATGAGCCCTAAGTTTGAAGTTTGTCTGAGAGTTCAATGCACTGGAATTA 1482
|||
Qy 346 AspAspSerAspLeuAlaILEPheILEAlaValILEILEuSerGLYAspArgProGLY 365
|||
Db 1483 GATGACATGACTGGCTATATTTATCTGCTGCTCATATTCTCAGGAGACCGCCCAAGGC 1542
|||
Qy 366 LeuLeuAsnValLYsProILEGLNAspILEGLNAspAsnLeuGLNAlaLeuGLNLeu 385
|||
Db 1543 TTGCTGAACGTGAACCCCATGACAGACATCCAGACAAACCTGCTCAGAGCCCTGGAAGCTG 1602
|||
Qy 386 GLNLeuLYsLeuAsnHISProGLNserSerGLNLeuPheAlaLYsLeuGLNLYsMet 405
|||
Db 1603 CAGCTCAAGCTGAATCACCCAGAGTCTCTCAGCTTTCGCCAAAGCTCTCCAGAAAGT 1662
|||
Qy 406 ThrAspLeuArgGLNILEValThrGLNHisValGLNLeuGLNValILELYsLYsThr 425
|||
Db 1663 ACAGACCTCAGGACAGATGCTCACAGACAGCAGTGCATGCTGATCAAGAAAGACAA 1722
|||
Qy 426 GLNThrAspMetSerLeuHISProLeuGLNGLNILEtyrLYsAspLeuTYR 443
|||
Db 1723 GAGACAGACATGAGCTTCACCCCTGCTCCAGAGATCTACAAAGACTTGTAT 1776
|||

```

```

; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 554
; LENGTH: 3301
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: n-a,t,g or c
US-09-954-456-554

Alignment Scores:
Pred. No.: 4,14e-119 Length: 3301
Score: 1101.50 Matches: 208
Percent Similarity: 80.98% Conservative: 56
Best Local Similarity: 63.80% Mismatches: 61
Query Match: 31.40% Indels: 1
DB: caps: 1

US-09-931-007a-1 (1-688) x US-09-954-456-554 (1-3301)
Qy 119 LysThrGLNLeuTYrAsnArgAsnLYsCYsGLNtyrCYsArgPheGLNLYsCYsLeuAla 138
|||
Db 683 AAGATTCAAGAGAAAGAACCGGCAACAGTGCAGTACTCCGCTTCCAAAGTGCCTGCA 742
|||
Qy 139 ValGLYMetSerHISAsnAlaILEArgPheGLYArgMetProGLNAlaGLNLYsGLYLYs 158
|||
Db 743 CTGGGCATGTCAACACAGCCTATCCGTTTGGTCGATGCGGAGGCTGAGAGAGAGAG 802
|||
Qy 159 LeuLeuAlaGLNILE---SerSerAspILEAspGLNLeuAsnProGLNserAspLeu 177
|||
Db 803 CTGCTGACAGGCGTCACTGCAACAGAGGAGGCCAGTACAAACCCACAGGTGGCGACCTG 862
|||
Qy 178 ArgAlaLeuAlaLYsHISLeuTYrAspSerTYrILELYsSerPheProLeuThrLYsAla 197
|||
Db 863 AAGGCTTCTCCAAAGCACATCTACATGCTTACCTGAAACCTTCACATGACCAAAAG 922
|||
Qy 198 LysAlaArgAlaILELeuThrGLYLYsThrThrAspLYsSerProPheValILEtyrAsp 217
|||
Db 923 AAGGCCCCAGCATCTCTACCGGCAAGCCAGCCACAGCGGCCCTTTGTATCCAGCAGC 982
|||
Qy 218 MetAsnSerLeuMetMetGLYGLNAspLYsILELYsPheLYsHISILEthrProLeuGLN 237
|||
Db 983 ATCGACACATTTGTGACGACAGAGAGAGAGGCGTGGTGGAAGACATGGTGAATGGCCTG 1042
|||
Qy 238 GLNGLNserLYsGLNValaILEArgILEPheGLNGLYCYsGLNpHeArgSerValaILE 257
|||
Db 1043 CTCCTCCATCAAGAGATAGGCTGCACGCTTCTTACCCCTCAGCCATGACCAACAGTGGAG 1102
|||
Qy 258 AlaValaGLNGLNILEthrGLNtyrAlaLYsSerILEProGLYpHeValAsnLeuAspLeu 277
|||
Db 1103 ACCGTCGGGAGCTCACAGTACGTTGCGCAAGAGCATCCCACTCTACAGAGCTCTTCTC 1162
|||
Qy 278 AsnAspGLNValaThrLeuLeuLYsTYrGLYValaHISGLNILEtyrThrMetLeuAla 297
|||
Db 1163 AACGACACAGGTACCTCTCTCAAGATAGCGTGCACGAGGCGCATCTTGGCATGCTGCC 1222
|||
Qy 298 SerLeuMetAsnLYsAspGLYValaLeuILEserGLNGLNGLNtyrGLNGLNpHeMetThrArgGLN 317
|||
Db 1223 TCTATCGTCAACAAAGAGGCGCTGCTGTGACCAAGCGCACTGCTTGTGTCACCCGTGAG 1282
|||

```

```
OY 318 PheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAla 337
      |||||.....|
Db 1283 TTCCTCGGAGGCGCTCGCAAAACCTTCAGTATCATTTAGAGCTTAAGTTGATTTGGT 1342

OY 338 ValLysPheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIle 357
      |||||.....|
Db 1343 GTCACATTCACAGCCCTCGAAGCTTGATGACAGGACCTGGCCCTATTCATTGGGGCCATC 1402

OY 358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluSplIleGlnAsp 377
      |||||.....|
Db 1403 ATTCTGTGGGAGACCGCGCCAGCTCATGACGTTCCACGGGTGGAGGCTATCCAGGAC 1462

OY 378 AsnLeuLeuGlnAlaLeuGluLeuGlnLysLeuAsnHisProGluSerGlnLeu 397
      |||||.....|
Db 1463 ACCATCTCGGAGCGCTCGCAATTCACATTCGACAGGCCACACCCCTGATGCCAGTACCTC 1522

OY 398 PheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGln 417
      |||||.....|
Db 1523 TTCGCCAAGCTGCTGCAAGATGGCTGACCTGGCGCACTGCTCACAGACAGCCCGCAG 1582

OY 418 LeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnIu 437
      |||||.....|
Db 1583 ATGATCGACAGCGATCAAGAGACCGCAACCGAGACTCGCTGCACCCCTCTCTCCAGGAG 1642

OY 438 IleTyrLysAspLeuTyr 443
      |||||.....|
Db 1643 ATCTACAAGACATGTAC 1660

RESULT 14
US-10-013-807-1
; Sequence 1, Application US/10013807
; Patent No. US20020112256A1
; GENERAL INFORMATION:
; APPLICANT: Guenther, Catherine
; APPLICANT: Phillips, Russell
; APPLICANT: Allen, Keith D.
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING
; TITLE OF INVENTION: PROLIFERATOR-ACTIVATED RECEPTOR GENE DISRUPTIONS
; FILE REFERENCE: R-683
; CURRENT APPLICATION NUMBER: US/10/013,807
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/254,916
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-013-807-1

Alignment Scores:
Pred. No.: 8,76e-117 Length: 1323
Score: 1076.50 Matches: 203
Percent Similarity: 80.06% Conservative: 58
Best Local Similarity: 62.27% Mismatches: 64
Query Match: 30.69% Indels: 1
DB: 12 Gaps: 1

US-09-931-007a-1 (1-688) x US-10-013-807-1 (1-1323)

OY 119 LysThrGlnLeuTyrAsnArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
      |||  |||  |||||.....|
Db 343 AAGATCCGAAAGAAAGACCGCAAGATGTCTAGTACTGCGGCTTCAGAGAGTCCCTGGCA 402

OY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGlnLys 158
      |||||.....|
Db 403 CTCGCGATGTCCGACAAACGCTATCCGCTTTGGACGATGCCGAGCGGAGAAAGAGAG 462

OY 159 LeuLeuAlaGluIle---SerSerAspIleAspGlnLeuAsnProGluSerLysAspLeu 177
      |||||.....|
Db 463 CTGGTGGGGGGGCTGACTGCGACGAGGGGGTCCAGACAAACCCCACTGGCGGAGCTG 522
```

```
OY 178 ArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAla 197
      |||||.....|
Db 523 AAGCGCTTCTCAGACACATCTACAAAGCGCTACCTGAAATACTTACACATGACCAAAAG 582

OY 198 LysAlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAsp 217
      |||||.....|
Db 583 AAGCGCGGAGGATCTCTCCCGCAAGTCCACCCACCAAGCGACCCCTTGTCTATCCAGCAG 642

OY 218 MetAsnSerLeuMetMetLysGluAspLysIleLysPheLysHisIleThrProLeuGln 237
      |||||.....|
Db 643 ATCGAGACACTGTGGCAGGCAAGAGCGCTGTGTGTGGAAACACTGTGTACAGCGGCTG 702

OY 238 GlnGlnSerLysValAlaIleArgIlePheGlnGlyGlnPheArgSerValGlu 257
      |||||.....|
Db 703 CGGCCCTACAGAGATATGATGATGACAGTGTTCTACCGCTGCCAGTCCACCACTGGAG 762

OY 258 AlaValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeu 277
      |||||.....|
Db 763 ACAGTCCGAGACTCACCGAGTTCCCAAGAACATCCCACTTCAGAGCGCTTCTCTC 822

OY 278 AsnAspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAla 297
      |||||.....|
Db 823 AATGACAGGTGACCTCTCAAGTATGGCTGCAAGAGCCATCTTGGCATGTGCTGGCC 882

OY 298 SerLeuMetAsnLysAspGlyValLeuLysSerGlnGlnGlyGlnPheMetThrArgLys 317
      |||||.....|
Db 883 TCCATCTGTAACAAAGAGCGGCTGCTGTGCGCCACAGCGCATGTGTCACCCAGCAG 942

OY 318 PheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAla 337
      |||||.....|
Db 943 TTCCTCGGAGGCTCGCAAGCCCTTCAGTACATCATTTGAGCCCAAGTTCGAGTTTCT 1002

OY 338 ValLysPheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIle 357
      |||||.....|
Db 1003 GTCAAGTTCAATGCCCTGAGCTGCATGACACTGAGCTGGCGCTTGTATGCGGCCATTC 1062

OY 358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAsp 377
      |||||.....|
Db 1063 ATTCTGTGGGAGACCGCGGCAAGCTCATGAAATGAGCCAGAGTGAAGCATCCAGGAC 1122

OY 378 AsnLeuLeuGlnAlaLeuGluLeuGlnLysLeuAsnHisProGluSerGlnLeu 397
      |||||.....|
Db 1123 ACCATCTCGGCGCTCGAAGATTCATCTGCGAGGTCACACCCCTGAGACCGAGTACCTC 1182

OY 398 PheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGln 417
      |||||.....|
Db 1183 TTCGCCAAGCTGCTCAGAGATGGCAGACCTGCGGCACTGTGTACTGAGATGCCAG 1242

OY 418 LeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnIu 437
      |||||.....|
Db 1243 ATGATCGACAGTGTCAAGAGCGAGAGTGAAGACTGTGTGCACCCCTGCTCCAGGAA 1302

OY 438 IleTyrLysAspLeuTyr 443
      |||||.....|
Db 1303 ATCTACAAGACATGTAC 1320

RESULT 15
US-10-033-528-1855
; Sequence 1855, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Mesgher, Madeleine Joy
; APPLICANT: Xu, Jiaqun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 1855
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 5
; OTHER INFORMATION: n = A,T,C or G
US-10-033-528-1855
```

Alignment Scores:

Pred. No.:	1.43e-73	Length:	428
Score:	702.00	Matches:	140
Percent Similarity:	99.29%	Conservative:	0
Best Local Similarity:	99.29%	Mismatches:	1
Query Match:	20.01%	Indels:	0
DB:	12	Gaps:	0

US-09-931-007a-1 (1-688) x US-10-033-528-1855 (1-428)

```
QY 246 ArgIlePheGlnGlyCysGlnPheArgSerValGluAlaValGlnIlePheGlnTyr 265
    |||||||
Db 6  CCATCTTTCAGGCTGCCAGTTGCTGCTGAGGCTGTGACAGATCAGAGATAT 65

QY 266 AlaIysSerIleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLys 285
    |||||||
Db 66  GCCAAAGCATTCCTGCTGTTGTAAATCTGACTGAACGACCAAGTAACCTCCTCAAA 125

QY 286 TyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyVal 305
    |||||||
Db 126  TATGGAGTCCAGAGATCATTTACCAATGCTGGCTCCTTGATGATTAAGATGGGCTT 185

QY 306 LeuIleSerGlnGlyGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysPro 325
    |||||||
Db 186  CTCATATCCGAGGCGCCAGAGCTTCATGACAAAGGAGTTCTAAAGAGCTGCGAAAGGCTT 245

QY 326 PheGlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeu 345
    |||||||
Db 246  TTTGGTGACTTATGAGCCCAAGTTGAGTTGCTGTGAAGTTCATGCACTGGAATTA 305

QY 346 AspAspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGly 365
    |||||||
Db 306  GATGACAGGAGCTTGCAATATTTATGCTGTCATTAATTCAGTGAGACCGCCAGGT 365

QY 366 LeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGluLeu 385
    |||||||
Db 366  TTGCTGAATGTAAGCCCATTTGAAGACATTCAGACAACCTGCCACAAGCCCTGGAGCTC 425

QY 386 Gln 386
    |||
Db 426  CAG 428
```

Search completed: February 25, 2003, 06:18:30
Job time : 293 secs

THIS PAGE BLANK (USPTO)